ROUGH PATHS AND 1D SDE WITH A TIME DEPENDENT DISTRIBUTIONAL DRIFT. APPLICATION TO POLYMERS.

François Delarue, Roland Diel

ROUGH PATHS AND 1D SDE WITH A TIME DEPENDENT DISTRIBUTIONAL DRIFT. APPLICATION TO POLYMERS.

FRANÇOIS DELARUE[1] AND ROLAND DIEL[2]

Laboratoire J.-A. Dieudonné,
Université de Nice Sophia-Antipolis and UMR CNRS 7351,
Parc Valrose, 06108 Nice Cedex 02, France.

Abstract. Motivated by the recent advances in the theory of stochastic partial differential equations involving nonlinear functions of distributions, like the Kardar-Parisi-Zhang (KPZ) equation, we reconsider the unique solvability of one-dimensional stochastic differential equations, the drift of which is a distribution, by means of rough paths theory. Existence and uniqueness are established in the weak sense when the drift reads as the derivative of a α-Hölder continuous function, $\alpha > 1/3$. Regularity of the drift part is investigated carefully and a related stochastic calculus is also proposed, which makes the structure of the solutions more explicit than within the earlier framework of Dirichlet processes.

1. Introduction

Given a family of continuous paths $(\mathbb{R} \ni x \mapsto Y_t(x))_{t\geqslant 0}$ with values in $\mathbb{R}$, we are interested in the solvability of the stochastic differential equation

$$\mathrm{d}X_t = \partial_x Y_t(X_t)\,\mathrm{d}t + \mathrm{d}B_t, \quad t \geqslant 0, \tag{1}$$

with a given initial condition, where $\partial_x Y_t$ is understood as the derivative of Y_t in the sense of distribution and $(B_t)_{t\geqslant 0}$ is a standard one-dimensional Wiener process.

When $\partial_x Y_t$ makes sense as a measurable function in $L^p_{\mathrm{loc}}(\mathbb{R}^d)$, for $p > 1$, pathwise existence and uniqueness are known to hold: See the earlier papers by Zvonkin [27] and Veretennikov [25] in the case when the derivative exists as a bounded function together with the more recent result by Krylov and Röckner [18] and the Saint-Flour Lecture Notes by Flandoli [8]. In the case when $\partial_x Y_t$ only exists as a distribution, existence and uniqueness have been only discussed within the restricted time homogeneous framework. When the field Y is independent of time, X indeed reads as a diffusion process with $\exp(-Y(x))\partial_x(\exp(Y(x))\partial_x)$ as generator. Then, solutions to (1) can be proved to be the sum of a Brownian motion and of a process of zero quadratic variation and are thus referred to as *Dirichlet* processes. In this setting, unique solvability can be proved to hold in the weak or strong sense according to the regularity of Y, see for example the papers by Flandoli, Russo and Wolf [9, 10] on the one hand and the paper by Bass and Chen [3] on the other hand.

In the current paper, we allow Y to depend upon time, making impossible any factorization of the generator of X under a divergence form and thus requiring a more systematic treatment of the singularity of the drift. In order to limit the technicality of the paper, the analysis is restricted to the case when the diffusion coefficient in (1) is 1, which is already, as explained

[1] delarue@unice.fr
[2] diel@unice.fr

right below, a really interesting case for practical purposes and which is, anyway, somewhat universal because of the time change property of the Brownian motion. As suggested in the aforementioned paper by Bass and Chen, pathwise existence and uniqueness are then no more expected to hold whenever the path Y_t has oscillations of Hölder type with a Hölder exponent strictly less than $1/2$. For that reason, we will investigate the unique solvability of (1) in the so-called weak sense by tackling a corresponding formulation of the martingale problem. Indeed, we will consider the case when Y_t is Hölder continuous, the Hölder exponent, denoted by α, being strictly greater than $1/3$, hence possibly strictly less than $1/2$, thus yielding solutions to (1) of weak type only, that is solutions that are not adapted to the underlying noise $(B_t)_{t\geqslant 0}$. At this stage of the introduction, it is worth mentioning that the threshold $1/3$ for the Hölder exponent of the path is exactly of the same nature as the one that occurs in the theory of rough paths.

Actually, the theory of rough paths will play a major role in our analysis. The strategy for solving (1) is indeed mainly inspired by the papers [27, 25, 18] we mentioned right above and consists in finding harmonic functions associated with the (formal) generator

$$\partial_t + \mathcal{L}_t := \partial_t + \frac{1}{2}\partial_x^2 + \partial_x Y_t(x)\partial_x. \tag{2}$$

Solving Partial Differential Equations (PDEs) driven by $\partial_t + \mathcal{L}_t$, say in the standard mild formulation, then requires to integrate with respect to $\partial_x Y_t(x)$ (in x), which is a non-classical thing. This is precisely the place where the rough paths theory initiated by Lyons (see [20, 19]) comes in: As recently exposed by Hairer in his seminal paper [14] on the KPZ equation, mild solutions to PDEs driven by $\partial_t + \mathcal{L}_t$ may be expanded as rough integrals involving the standard heat kernel on the one hand and the 'rough' increments $\partial_x Y_t$ on the other hand. In our case, we are interested in the solutions of the PDE

$$\partial_t u_t(x) + \mathcal{L}_t u_t(x) = f_t(x), \tag{3}$$

when set on a cylinder of the form $[0,T]\times\mathbb{R}$ (with a terminal boundary condition at time T) and when driven by a smooth function f. Solutions obtained by letting the source term f vary generates a large enough 'core' in order to apply the standard martingale problem approach by Stroock and Varadhan [23] and thus to characterize the laws of the solutions to (1).

Unfortunately, although such a strategy seems quite clear, some precaution is in fact needed. When α is between $1/3$ and $1/2$, which is the typical range of application of Lyons' theory, the expansion of mild solutions as rough integrals involving the heat kernel and the increments of $\partial_x Y_t$ is not so straightforward. It is indeed not enough to assume that the path $\mathbb{R}\ni x\mapsto Y_t(x)$ has a rough path structure for any given time $t\geqslant 0$. As explained in detail in Section 2, the rough path structures, when taken at different times, also interact, asking for the existence, at any time $t\geqslant 0$, of a 'lifted' 2-dimensional rough path with Y_t as first coordinate. We refrain from detailing the shape of such a lifting right here as it is longly discussed in the next section. We just mention that, in Hairer [14], the whole family $(Y_t(x)))_{t\geqslant 0,x\in\mathbb{R}}$ has a Gaussian structure, which permits to construct the lifting by means of generic results on rough paths for Gaussian processes, see Friz and Victoir [12]. Existence of the lifting under more general assumptions is a thus a challenging question, which is (partially) addressed in Section 5: The lifting is proved to exist in several cases, including that when $\alpha > 1/2$, when $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ has some smoothness in time (and in particular when

it is time homogeneous) or when it satisfies a forward SPDE driven by a space-time white noise and by a kernel with the same kind of singularities as the Gaussian kernel. Another difficulty is that, contrary to Hairer [14] in which the problem is set on the torus, the PDE is here set on a non-compact domain. This requires an additional analysis of the growth of the solutions in terms of the behavior of $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ for large values of $|x|$, such an analysis being essential for discussing the non-explosion of the solutions to (1).

Besides existence and uniqueness, it is also of great interest to understand the specific dynamics of the solutions to (1). Part of the paper is thus dedicated to a careful analysis of the infinitesimal variation of X, that is of the asymptotic behavior of $X_{t+h} - X_t$ as h tends to 0. In this perspective, we prove that the increments of X may be split into two pieces: a Brownian increment as suggested by the initial writing of Eq. (1) and a sort of drift term, the magnitude of which is of order $h^{(1+\beta)/2}$, for some $\beta > 0$ that is nearly equal to α. Such a decomposition is much stronger than the standard decomposition of a Dirichlet process into the sum of a martingale and of a zero quadratic variation process. Somehow it generalizes the one obtained by Bass and Chen in the time homogeneous framework when $\alpha \geqslant 1/2$. As a typical example, $(1+\beta)/2$ is nearly equal to $3/4$ when Y_t is almost $1/2$-Hölder continuous, which fits for instance the framework investigated by Hairer [14]. In particular, except trivial cases when the distribution is a true function, integration with respect to the drift term in (1) cannot be performed as a classical integration with respect to a function of bounded variation. In fact, since the value of $(1+\beta)/2$ is strictly larger than $1/2$, it makes sense to understand the integration with respect to the drift term as a kind of Young integral, on the same model as the one developed by Young in the earlier paper [26]. We here say 'a kind of Young integral' and not 'a Young integral' directly since, as we will see in the analysis, it sounds useful to develop a stochastic version of Young's integration, that is a Young-like integration that takes into account the probabilistic notion of adaptedness as it is the case in Itô's calculus.

In the end, we prove that, under appropriate assumptions on the regularity of the field $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$, Eq. (1) is uniquely solvable in the weak sense (for a given initial condition) and that the solution reads as

$$\mathrm{d}X_t = b(t, X_t, \mathrm{d}t) + \mathrm{d}B_t, \tag{4}$$

where $b(t,x,h)$ is a function from $[0,+\infty)\times\mathbb{R}\times[0,+\infty)$ into $\mathbb{R}$ and the integral with respect to $b(t, X_t, \mathrm{d}t)$ makes sense as a stochastic Young integral, the magnitude of $b(t, X_t, \mathrm{d}t)$ being of order $\mathrm{d}t^{(1+\beta)/2}$.

The examples we have in mind are twofold. The first one is the so-called 'Brownian motion in a time-dependent random environment' or 'Brownian motion in a time-dependent random potential'. Indeed, much has been said about the long time behavior of the Brownian motion in a time-independent random potential such as the Brownian motion in a Brownian potential, see for example [2, 5, 6, 15, 16, 22, 24]. We expect our paper to be a first step forward toward a more general analysis of one-dimensional diffusions in a time-dependent random potential, even if, in the current paper, nothing is said about the long run behavior of the solutions to (1), this question being left to further investigations. As already announced, the second example we have in mind is the so-called Kardar-Parisi-Zhang (KPZ) equation (see [17]), to which much attention has been paid recently, see among others the seminal papers by Bertini and Giacomin [4] and Hairer [14] about the well-posedness on the one hand and by Amir, Corwin and Quastel [1] about the long time behavior on the other hand.

In this framework, Y must be thought as a realization of the time-reversed solution of the KPZ equation, that is $Y_t(x) = u(\omega, T-t, x)$, T being positive and $u(\omega,\cdot,\cdot)$ denoting the random solution to the KPZ equation and being defined either as in Bertini and Giacomin by means of the Cole-Hopf transform or as in Hairer by means of renormalization arguments. Then, it is worth noting that, in this framework, Eq. (1) reads as the equation for describing the dynamics of the canonical path $(w_t)_{0\leqslant t\leqslant T}$ on the canonical space $\mathcal{C}([0,T],\mathbb{R})$ under the polymer measure

$$\exp\left(\int_0^T \dot\zeta(t,w_t)\,\mathrm{d}t\right)\mathrm{d}\mathbb{P}(w),$$

where $\dot\zeta$ is a space-time white noise and $\mathbb{P}$ is the Wiener measure, the white noise being independent of the realizations of the Wiener process under $\mathbb{P}$. In this perspective, our result provides a *quenched* description of the infinitesimal dynamics of the polymer. As for the analysis of one-dimensional processes in a random potential, additional results about the long time behavior would be of great interest. Again, we hope to handle this question in future works.

The paper is organized as follows. We remind the reader of the rough paths theory in Section 2. Main results about the solvability of (1) are also exposed in Section 2. Section 3 is devoted to the analysis of partial differential equations driven by the operator (2). In Section 4, we propose a stochastic variant of Young's integral in order to give a rigorous meaning to (4). We discuss in Section 5 the construction of the 'rough' iterated integral that makes the whole construction work. Finally, in Section 6, we explain the connection with the KPZ equation.

2. General Strategy and Main Results

Our basic strategy to define a solution to the SDE (1) relies on a suitable adaptation of Zvonkin's method for solving SDEs driven by a bounded and measurable drift (see [27]) and of Stroock and Varadhan's martingale problem (see [23]). The main point is to transform the original equation into a martingale. For sure such a strategy requires a suitable version of Itô's formula and henceforth a right notion of harmonic functions for the generator of the diffusion process (1). This is precisely the point where the rough paths theory comes in, on the same model as it does in the paper by Hairer for solving the KPZ equation.

This section is thus devoted to a sketchy presentation of rough paths theory and then to an appropriate reformulation of Zvonkin's method.

2.1. Rough paths on a segment.

In order to introduce elements of rough path theory, we will use the approach due to Gubinelli in [13].

Given $\alpha\in(0,1]$ and $n\in\mathbb{N}\backslash\{0\}$ and given a segment $\mathbb{I}\subset\mathbb{R}$, we denote by $\mathcal{C}^\alpha(\mathbb{I},\mathbb{R}^n)$ the set of α-Hölder continuous functions f from $\mathbb{I}$ to $\mathbb{R}^n$. We then define the seminorm

$$\|f\|_\alpha^{\mathbb{I}} := \sup_{x,y\in\mathbb{I},x\neq y}\frac{|f(y)-f(x)|}{|y-x|^\alpha} \text{ and the norm } [\![f]\!]_\alpha^{\mathbb{I}} := \|f\|_\infty^{\mathbb{I}} + (1\vee\max_{x\in\mathbb{I}}|x|)^{-\frac{\alpha}{2}}\|f\|_\alpha^{\mathbb{I}},$$

with $\|f\|_\infty^{\mathbb{I}} := \sup_{x\in\mathbb{I}}|f(x)|$ and $a\vee b=\max(a,b)$. Note that the factor $(1\vee\max_{x\in\mathbb{I}}|x|)^{-\alpha/2}$ is somewhat useless and could be replaced by 1 at this stage of the paper. Actually it will really matter in the sequel, when considering paths over the whole line. Similarly, we denote by $\mathcal{C}_2^\alpha(\mathbb{I},\mathbb{R}^n)$ the set of functions $\mathscr{R}$ from $\mathbb{I}^2$ to $\mathbb{R}^n$ such that $\mathscr{R}(x,x)=0$ for every x and with

finite norm $\|\mathscr{R}\|_\alpha^{\mathbb{I}} := \sup_{x,y\in\mathbb{I},x\neq y}\{|\mathscr{R}(x,y)|/|y-x|^\alpha\}$. (Functionals defined on the product space $\mathbb{R}^2$ will be denoted by calligraphic letters).

For $\alpha \in (1/3, 1]$, we call α-rough path (on $\mathbb{I}$) a pair $(W, \mathscr{W})$ where $W \in \mathcal{C}^\alpha(\mathbb{I}, \mathbb{R}^n)$ and $\mathscr{W} \in \mathcal{C}_2^{2\alpha}(\mathbb{I}, \mathbb{R}^{n^2})$ such that, for any indices $i, j \in \{1, \dots, n\}$, the following relation holds:

$$\mathscr{W}^{i,j}(x,z) - \mathscr{W}^{ij}(x,y) - \mathscr{W}^{ij}(y,z) = (W^i(y) - W^i(x))(W^j(z) - W^j(y)), \quad x \leqslant y \leqslant z. \tag{5}$$

We then denote by $\mathcal{R}^\alpha(\mathbb{I}, \mathbb{R}^n)$ the set of α-rough paths; with a slight abuse of notation, we will often only write W for the rough path $(W, \mathscr{W})$. The quantity $\mathscr{W}^{i,j}(x,y)$ must be understood as a value for the iterated integral (or cross integral) "$\int_x^y (W^i(z) - W^i(x))\,\mathrm{d}W^j(z)$" of W with respect to itself (we will also use the tensorial product "$\int_x^y (W(z) - W(x)) \otimes \mathrm{d}W(z)$" to denote the product between coordinates). Whenever $\alpha = 1$, such an integral exists in a standard sense. Whenever $\alpha > 1/2$, it exists as well, but in the so-called Young's sense (see [26, 19]). Whenever $\alpha \in (1/3, 1/2]$, which is the typical range of values in rough paths theory, there is no more a canonical way to define the cross integral and it must be given a priori in order to define a proper integration theory with respect to $\mathrm{d}W$. In that framework, condition (5) imposes some consistency in the behavior of $\mathscr{W}$ when intervals of integration are concatenated. Of course, $\mathscr{W}$ plays a role in the range $(1/3, 1/2]$ only, but in order to avoid any distinction between the cases $\alpha \in (1/3, 1/2]$ and $\alpha \in (1/2, 1]$, we will refer to the pair $(W, \mathscr{W})$ in both cases, even when $\alpha > 1/2$, in which case $\mathscr{W}$ will be just given by the iterated integral of W.

Given $W \in \mathcal{R}^\alpha(\mathbb{I}, \mathbb{R}^n)$ as above, the point is then to define the integral "$\int_x^y v(z)\,\mathrm{d}W(z)$" of some function v (from $\mathbb{I}$ into itself) with respect to the coordinates of $\mathrm{d}W$ for some $[x, y] \subset \mathbb{I}$. When v belongs to $\mathcal{C}^\beta(\mathbb{I}, \mathbb{R})$, for $\beta > 1 - \alpha$, Young's theory applies, without any further reference to the second-order structure $\mathscr{W}$ of W. Whenever $\beta \leqslant 1 - \alpha$, Young's theory fails, but, in the typical example when v is $W - W(x)$ itself (or one coordinate of $W - W(x)$), the integral is well-defined as it is precisely given by $\mathscr{W}$. In order to benefit from the second-order structure of $\mathscr{W}$ for integrating a more general v, the increments of v must actually be structured in a similar fashion to that of W. This motivates the following notion: For $\beta \in (1/3, 1 - \alpha]$, we say that a path v is β-controlled by W if $v \in \mathcal{C}^\beta(\mathbb{I}, \mathbb{R})$ and there is a function $\partial_W v \in \mathcal{C}^\beta(\mathbb{I}, \mathbb{R}^n)$ such that the remainder term

$$\mathscr{R}^v(x,y) := v(y) - v(x) - \partial_W v(x)(W(y) - W(x)), \quad x, y \in \mathbb{I},$$

is in $\mathcal{C}_2^{2\beta}(\mathbb{I}, \mathbb{R})$ (pay attention that, in the above formula, $\partial_W v(x)$ reads as a row vector -as it is often the case for gradients- and $(W(y) - W(x))$ as a column vector). For $\beta \in (1 - \alpha, 1]$, this notion is pretty useless: We then say that a path v is β-controlled by W if v is simply in $\mathcal{C}^\beta(\mathbb{I}, \mathbb{R})$, which is to say that the above holds with $\partial_W v(x) = 0$ and $\mathscr{R}^v(x,y) = v(y) - v(x)$. For $\beta \in (1/3, 1]$, we denote by $\mathcal{B}^\beta(\mathbb{I}, W)$ the set of such pairs $(v, \partial_W v)$. We emphasize that $\partial_W v$ may not be uniquely defined, but, when there is no possible confusion on the value of $\partial_W v$, we will only write v for $(v, \partial_W v)$.

We are then able to define the integral of a function v that is controlled by W (see [14, 13]):

Theorem 1. *Given $\alpha, \beta \in (1/3, 1]$, let $W \in \mathcal{R}^\alpha(\mathbb{I}, \mathbb{R}^n)$ be a rough path and $v \in \mathcal{B}^\beta(\mathbb{I}, W)$ be a path controlled by W. For two reals $x < y$ in $\mathbb{I}$, consider the compensated (vectorial)*

Riemann sum:

$$S(\Delta) := \sum_{i=0}^{N-1} \Big\{ v(x_i)\big(W(x_{i+1}) - W(x_i)\big) + \partial_W v(x_i) \mathscr{W}(x_i, x_{i+1}) \Big\}$$

where $\Delta = (x = x_0 < \cdots < x_N = y)$ is a partition of $[x, y]$ (above $\partial_W v(x_i)$ reads as a row vector and $\mathscr{W}(x_i, x_{i+1})$ as a matrix). Then, as the step size $\pi(\Delta)$ of the partition converges to zero, $S(\Delta)$ converges to a limit denoted by $\int_x^y v(z)\,\mathrm{d}W(z)$, the value of which is independent of the choice of the approximating partitions. Moreover, there exists a constant $C = C(n, \alpha, \beta)$ such that

$$\begin{aligned} &\left| \int_x^y v(z)\,\mathrm{d}W(z) - v(x)\big(W(y) - W(x)\big) - \partial_W v(x) \mathscr{W}(x, y) \right| \\ &\qquad \leqslant C \left(\|\mathscr{W}\|_{2\alpha}^{[x,y]} \|\partial_W v\|_{\beta}^{[x,y]} |y - x|^{2\alpha+\beta} + \|W\|_{\alpha}^{[x,y]} \|\mathscr{R}^v\|_{2\beta}^{[x,y]} |y - x|^{\alpha+2\beta} \right). \end{aligned}$$

Observe in particular that, with our prescribed range of values for α and β, the exponents $2\alpha+\beta$ and $\alpha+2\beta$ are (strictly) greater than 1, thus making the right hand side much smaller than the length of the interval $[x, y]$. It is worth mentioning that this observation is crucial for proving the convergence of $S(\Delta)$ as the step size tends to 0.

Stability of the integral with respect to W is a crucial question for practical purposes. In particular, it is really sound to wonder about the stability of the integral by regularization. Replacing (v, W) by a sequence of smooth approximations $(v^n, W^n)_{n\geqslant 1}$, the question is to decide whether the (classical) integrals of the $(v^n)_{n\geqslant 1}$'s with respect to the approximated paths are indeed close to the rough integral of v with respect to W. As well-guessed, the answer turns out to be false in full generality, as it would provide a canonical construction of the integral if it were true. Actually, it turns out to be true if the convergence holds in the rough paths sense, that is $[\![W - W^n]\!]_{\alpha}^{\mathbb{I}} + \|\mathscr{W} - \mathscr{W}^n\|_{2\alpha}^{\mathbb{I}}$ also tends to 0 as n tends to the infinity ($\mathscr{W}^n$ standing for the true iterated integral of W^n), in which case we say that the rough path W (or $(W, \mathscr{W})$) is geometric, and $[\![v - v^n]\!]_{\beta}^{\mathbb{I}} + [\![\partial_W v - \partial_{W^n} v^n]\!]_{\beta}^{\mathbb{I}} + [\![\mathscr{R}^v - \mathscr{R}^{v^n}]\!]_{2\beta}^{\mathbb{I}}$ also tends to 0 as n tends to the infinity.

2.2. **Time indexed families of rough paths.** It is well-guessed that, in order to handle (1), we have in mind to choose $W(x) = Y_t(x)$, $x \in \mathbb{R}$, and to apply rough paths theory at any fixed time $t \geqslant 0$ (thus requiring to choose $\mathbb{I} = \mathbb{R}$ and subsequently to extend the notion of rough paths to the whole $\mathbb{R}$, which will be done in the next paragraph). Anyhow a difficult aspect for handling (1) is precisely that $(Y_t(x))_{t\geqslant 0, x\in\mathbb{R}}$ is time dependent. If it were time homogeneous, part of the analysis we provide here would be useless: we refer for instance to [9, 10, 3]. From the technical point of view, the reason is that, in the homogeneous framework, the analysis of the generator of the process X reduces to the analysis of a standard one-dimensional ordinary differential equation. Whenever coefficients depend on time, the connection with ODEs boils down, thus asking for non-trivial refinements. From the intuitive point of view, time-inhomogeneity makes things much more challenging as the underlying differential structure in space varies at any time: In order to integrate with respect to $\partial_x Y_t(x)$ in the rough paths sense, the second-order structure of the rough paths must be defined first and it is well-understood that it is then time-dependent as well. This says that the problem consists of a time-indexed family of rough paths, but, a priori (and unfortunately), it is by no means clear whether defining the rough paths time by time can be enough for handling

the problem. Actually, as we explain right below, it may not be enough as the rough paths structures actually interact one with the others, thus requiring an additional assumption on $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$.

As in the previous subsection, we first limit our exposition of time-dependent rough paths to the case when x lives in a segment $\mathbb{I}$. For some time horizon $T>0$, and for $\alpha,\gamma>0$, we define the following (semi-)norms for continuous functions $f:[0,T)\times\mathbb{I}\to\mathbb{R}^n$ and $\mathscr{M}:[0,T)\times\mathbb{I}^2\to\mathbb{R}^n$:

$$\|f\|_{\gamma,\alpha}^{[0,T)\times\mathbb{I}}:=\sup_{\substack{x,y\in\mathbb{I},x\neq y,\\ 0\leqslant s<t<T}}\frac{|f_t(y)-f_s(x)|}{|t-s|^\gamma+|y-x|^\alpha}\quad\text{and}\quad\|\mathscr{M}\|_{0,\alpha}^{[0,T)\times\mathbb{I}}:=\sup_{\substack{x,y\in\mathbb{I},x\neq y\\ 0\leqslant t<T}}\frac{|\mathscr{M}(t,x,y)|}{|y-x|^\alpha},$$

together with

$$[\![f]\!]_{\gamma,\alpha}^{[0,T)\times\mathbb{I}}:=\|f\|_\infty^{[0,T)\times\mathbb{I}}+(1\vee\max_{x\in\mathbb{I}}|x|)^{-\frac{\alpha}{2}}\|f\|_{\gamma,\alpha}^{[0,T)\times\mathbb{I}}.$$

We then define the spaces $\mathcal{C}^{\gamma,\alpha}([0,T)\times\mathbb{I},\mathbb{R}^n)$ and $\mathcal{C}_2^{\gamma,\alpha}([0,T)\times\mathbb{I},\mathbb{R}^n)$ accordingly.

For $\alpha\in(1/3,1/2]$, we call time dependent α-rough path a pair function $(W_t,\mathscr{W}_t)_{0\leqslant t<T}$ where $W\in\mathcal{C}([0,T)\times\mathbb{I},\mathbb{R}^n)$ and $\mathscr{W}\in\mathcal{C}([0,T)\times\mathbb{I}^2,\mathbb{R}^{n^2})$ such that, for any $t\in[0,T)$, the pair $(W_t,\mathscr{W}_t)$ is an α-rough path and

$$\|(W,\mathscr{W})\|_{0,\alpha}^{[0,T)\times\mathbb{I}}:=\sup_{t\in[0,T)}\left\{\|W_t\|_\alpha^{\mathbb{I}}+\|\mathscr{W}_t\|_{2\alpha}^{\mathbb{I}}\right\}<\infty. \tag{6}$$

We denote by $\mathcal{R}^\alpha([0,T)\times\mathbb{I},\mathbb{R}^n)$ the set of time-dependent α-rough paths endowed with the seminorm $\|\cdot\|_{0,\alpha}^{[0,T)\times\mathbb{I}}$. For $\beta\in(1/3,1-\alpha]$, we then say that $v\in\mathcal{C}([0,T)\times\mathbb{I},\mathbb{R})$ is β-controlled by the paths $(W_t)_{0\leqslant t<T}$ if $v\in\mathcal{C}^{\beta/2,\beta}([0,T)\times\mathbb{I},\mathbb{R})$ and there exists a function $\partial_W v\in\mathcal{C}^{\beta/2,\beta}([0,T)\times\mathbb{I},\mathbb{R}^n)$ such that, for any $t\in[0,T)$, the remainder term

$$\mathscr{R}^{v_t}(x,y):=v_t(y)-v_t(x)-\partial_W v_t(x)(W_t(y)-W_t(x)) \tag{7}$$

is in $\mathcal{C}_2^{2\beta}(\mathbb{I},\mathbb{R}^n)$. For $\beta\in(1-\alpha,1]$, we always say that $v\in\mathcal{C}^{\beta/2,\beta}([0,T)\times\mathbb{I},\mathbb{R}^n)$ is β-controlled by the paths $(W_t)_{0\leqslant t<T}$. In that case, $\partial_W v=0$ and $\mathscr{R}^v(t,x,y)=v_t(y)-v_t(x)$. For any $\beta\in(1/3,1]$, we then denote by $\mathcal{B}^\beta([0,T)\times\mathbb{I},(W_t)_{0\leqslant t<T})$ (or simply by $\mathcal{B}^\beta([0,T)\times\mathbb{I},W)$) the set of pairs $(v,\partial_W v)$.

2.3. **Rough paths on the whole line.** So far, we have only defined rough paths (or time dependent rough paths) on segments. As Eq. (1) is set on the whole $\mathbb{R}$, we must extend the definition from segments to $\mathbb{R}$. As well-guessed, the point is to specify the behavior at infinity of the underlying (rough) paths and of the corresponding functions that are controlled by these paths.

In the case when the family $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ is differentiable in x, the typical assumption for solving (1) (and in particular to prevent any blow-up) consists in requiring $(\partial_x Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ to be at most of linear growth in x. In our setting, $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ is singular and it makes no sense to discuss the growth of its derivative. The point is thus to control the growth of the local Hölder norm of $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ together with (as shown later) the growth of the local Hölder norm of the associated iterated integral.

This motivates the following definition. For $\alpha\in(1/3,1]$ and $\chi>0$, we call α-rough path (on $\mathbb{R}$) with rate χ a pair $(W,\mathscr{W})$ such that, for any $r\geqslant 1$, the restriction of $(W,\mathscr{W})$ to

$[-r,r]$ is in $\mathcal{R}^\alpha([-r,r])$, and

$$\kappa_{\alpha,\chi}\big(W,\mathscr{W}\big) := \sup_{r\geqslant 1} \frac{\|W\|_\alpha^{[-r,r]}}{r^\chi} + \frac{\|\mathscr{W}\|_{2\alpha}^{[-r,r]}}{r^{2\chi}} < \infty. \tag{8}$$

We denote by $\mathcal{R}^{\alpha,\chi}(\mathbb{R},\mathbb{R}^n)$ the set of all such $(W,\mathscr{W})$.

This definition extends to time-dependent families of rough paths. Given $T>0$, we say that $(W_t,\mathscr{W}_t)_{0\leqslant t<T}$ belongs to $\mathcal{R}^{\alpha,\chi}([0,T)\times\mathbb{R},\mathbb{R}^n)$ if

$$\kappa_{\alpha,\chi}\big((W_t,\mathscr{W}_t)_{0\leqslant t<T}\big) := \sup_{t\in[0,T)}\sup_{r\geqslant 1} \frac{\|W\|_\alpha^{[0,T)\times[-r,r]}}{r^\chi} + \frac{\|\mathscr{W}\|_{2\alpha}^{[0,T)\times[-r,r]}}{r^{2\chi}} < \infty. \tag{9}$$

In a similar way, we must specify the admissible growth of the functions that are controlled by rough paths on the whole real line. As shown later, a quite comfortable framework is then to require exponential bounds. Given $(W,\mathscr{W})\in\mathcal{R}^{\alpha,\chi}(\mathbb{R},\mathbb{R}^n)$ and $\vartheta\geqslant 1$, we thus say that a function $v:\mathbb{R}\to\mathbb{R}$ is in $\mathcal{B}^{\beta,\vartheta}(\mathbb{R},W)$ for some $\beta\in(1/3,1]$ if, for any segment $\mathbb{I}\subset\mathbb{R}$, the restriction of v to $\mathbb{I}$ is in $\mathcal{B}^\beta(\mathbb{I},W)$ and

$$\Theta^\vartheta(v) := \sup_{r\geqslant 1}\Big[e^{-\vartheta r}\Big(\llbracket v\rrbracket_\beta^{[-r,r]} + \tfrac{1}{2}\llbracket\partial_W v\rrbracket_\beta^{[-r,r]} + r^{-\beta}\|\mathscr{R}^v\|_{2\beta}^{[-r,r]}\Big)\Big] < \infty. \tag{10}$$

(With an abuse of notation we omit to specify the dependence upon $\partial_W v$ in $\Theta^\vartheta(v)$.) Similarly, given $(W_t,\mathscr{W}_t)_{0\leqslant t<T}\in\mathcal{R}^{\alpha,\chi}([0,T)\times\mathbb{R},\mathbb{R}^n)$, we say that a function $v:[0,T)\times\mathbb{R}\to\mathbb{R}$ is in $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$ if the restriction of v to $[0,T)\times[-r,r]$ is in $\mathcal{B}^\beta([0,T)\times[-r,r],W)$ for any $r\geqslant 1$ and, for some $\lambda\geqslant 0$,

$$\begin{aligned}&\Theta_T^{\vartheta,\lambda}(v)\\ &\quad:= \sup_{\substack{r\geqslant 1\\ t\in[0,T)}}\Big[\frac{1}{E_T^{\vartheta,\lambda}(t,r)}\Big(\llbracket v\rrbracket_{\beta/2,\beta}^{[t,T)\times[-r,r]} + \tfrac{1}{2}\llbracket\partial_W v\rrbracket_{\beta/2,\beta}^{[t,T)\times[-r,r]} + \big((T-t)^{\frac{\beta}{2}}+r^{-\beta}\big)\|\mathscr{R}^{v_t}\|_{2\beta}^{[-r,r]}\Big)\Big]\end{aligned}$$

is finite, where $E_T^{\vartheta,\lambda}(t,r) := \exp[\lambda(T-t)+\vartheta r(1+T-t)]$. Note that the set $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$ doesn't depend on λ, but that $\Theta_T^{\vartheta,\lambda}(v)$ does. The reason why we consider $[t,T)$ and not $[0,t]$ in the above bound follows from the standard connection between stochastic differential equations (or more generally Markov processes) and backward partial differential equations, which is exactly the subject of the next paragraph. Put it differently, exponential growth propagates in a backward direction in the analysis of (3).

By Theorem 1, we can easily obtain a control of the integral $\int v_t\,\mathrm{d}Y_t$ by the norm $\Theta_T^{\vartheta,\lambda}(v)$:

Lemma 2. *There exists a constant $C=C(n,\alpha,\beta)$, such that for any $\vartheta,\lambda,r\geqslant 1$, any $v\in\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$ and for any $(t,x,y)\in[0,T)\times[-r,r]^2$,*

$$\begin{aligned}&\Big|\int_x^y \big(v_t(z)-v_t(x)\big)\,\mathrm{d}W_t(z)\Big| \leqslant C\kappa_{\alpha,\chi}\big(W_t,\mathscr{W}_t\big)\Theta_T^{\vartheta,\lambda}(v)E_T^{\vartheta,\lambda}(t,r)\times\mathscr{D}(t,r,y-x)\\ &\Big|\int_x^y v_t(z)\,\mathrm{d}W_t(z)\Big| \leqslant C\kappa_{\alpha,\chi}\big(W_t,\mathscr{W}_t\big)\Theta_T^{\vartheta,\lambda}(v)E_T^{\vartheta,\lambda}(t,r)\times\big[|y-x|^\alpha r^\chi+\mathscr{D}(t,r,y-x)\big],\end{aligned}$$

with $\mathscr{D}(t,r,z) := |z|^{2\alpha}r^{2\chi}+|z|^{2\alpha+\beta}r^{2\chi+\frac{\beta}{2}}+|z|^{\alpha+2\beta}r^\chi(r^\beta+(T-t)^{-\frac{\beta}{2}})$.

2.4. **Enlargement of the rough path structure.** As we discussed right above, it is quite crucial to understand how the time dependent rough path structures of the drift $(Y_t(x))_{t\geqslant 0, x\in\mathbb{R}}$ interact one with the others as time varies.

Formally the generator associated with (1) reads $\mathcal{L} = \partial_t + \partial_x(Y_t(x))\partial_x + (1/2)\partial^2_{xx}$. This suggests that, on $[0,T)\times\mathbb{R}$, harmonic functions (that is zeros of the generator) read as

$$u_t(x) = \int_{\mathbb{R}} p_{T-t}(x-z)u_T(z)\,\mathrm{d}z + \int_t^T \int_{\mathbb{R}} p_{r-t}(x-z)\partial_x u_r(z)\,\mathrm{d}Y_r(z)\,\mathrm{d}r, \quad x\in\mathbb{R},$$

where p denotes the standard heat kernel. In the case when the boundary condition of the function v is given by $u_T(x) = x$, a formal expansion for $\partial_x u_t(x)$ in the neighborhood of T gives

$$\begin{aligned}\partial_x u_t(x) \sim 1 &+ \int_t^T \int_{\mathbb{R}} \partial_x p_{r-t}(x-z)\,\mathrm{d}Y_r(z)\,\mathrm{d}r \\ &+ \int_t^T \int_{\mathbb{R}} \partial_x p_{r-t}(x-z)\left\{\int_r^T \int_{\mathbb{R}} \partial_x p_{s-r}(z-u)\,\mathrm{d}Y_s(u)\,\mathrm{d}s\right\}\mathrm{d}Y_r(z)\,\mathrm{d}r + \dots\end{aligned}$$

In the first-order term of the expansion, the space integral makes sense as the singularity can be transferred from Y_r onto $\partial_x p_{r-t}(x-z)$, provided the integration by parts is licit: Using the approximation argument discussed above, it is indeed licit when the rough path is geometric. In order to give a sense to this first-order term, the point is then to check that the resulting singularity in time is integrable: this question is addressed in Section 3. Unfortunately, the story is much less simple for the second order term. Indeed, any formal integration by parts leads to a term involving a 'cross' integral between the spatial increments of Y, but taken at different times... This is exactly the place where rough path structures, considered at different times, interact.

We refrain from detailing the computations as this stage of the paper and feel more convenient to reject their presentation to Section 3 below. Basically, the point is to give, at any time $t\in[0,T)$, a sense to the integral $\int_x^y Z_t^T(z)\,\mathrm{d}Y_t(z)$, where

$$\forall t\in[0,T),\ \forall x\in\mathbb{R},\ Z_t^T(x) = \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{r-t}(x-z)(Y_r(z)-Y_r(x))\,\mathrm{d}z\,\mathrm{d}r. \tag{11}$$

Assuming that $\sup_{0\leqslant t<T}\sup_{x,y\in\mathbb{R}}[(1+|x|^\chi+|y|^\chi)^{-1}[\![Y_t]\!]_\alpha^{[x,y]}]$ is finite (for some $\chi>0$), the above integral is well-defined (thanks to standard Gaussian estimates, see Section 3). In order to make sure that the cross integral of Z_t^T with respect to Y_t exists, the point is to assume that the pair (Y_t, Z_t^T) can be lifted up to a rough path of dimension 2, which is to say that there exists some $\mathscr{W}^T$ with values in $\mathbb{R}^4$ such that $((Y,Z^T),\mathscr{W}^T)$ is an α-time dependent rough path, for some $\alpha>1/3$. We will see in Section 5 conditions under which such a lifting $\mathscr{W}^T$ indeed exists.

2.5. **Generator of the diffusion and related Dirichlet problem.** We now provide some solvability results for the Dirichlet problem driven by the operator $\partial_t + \partial_x Y(t,\cdot)\partial_x + (1/2)\partial^2_{xx}$ on $[0,T]\times\mathbb{R}$, for some $T>0$.

Definition 3. *Given $Y\in\mathcal{C}([0,T)\times\mathbb{R},\mathbb{R})$, assume that there exists $\mathscr{W}^T$ such that $(W^T = (Y,Z^T),\mathscr{W}^T)$ belongs to $\mathcal{R}^{\alpha,\chi}([0,T)\times\mathbb{R},\mathbb{R}^2)$ with $\alpha>1/3$ and $\chi<\beta/2$. Given an exponent $\vartheta\geqslant 1$, a real $\beta\in(1/3,\alpha)$ and a function $f\in L^\infty((0,T)\times\mathbb{R})$, we say that a continuous*

function $u : [0,T] \times \mathbb{R} \to \mathbb{R}$, continuously differentiable with respect to x, such that the restriction of $\partial_x u$ to $[0,T) \times \mathbb{R}$ is in $\mathcal{B}^{\beta,\vartheta}([0,T) \times \mathbb{R}, W^T)$, is a mild solution on $[0,T] \times \mathbb{R}$ to the problem $\mathcal{P}(Y,f,T)$:

$$\mathcal{L}v = f, \quad \text{with} \quad \mathcal{L}v := \partial_t v + \mathcal{L}_t v,$$

if, for any $(t,x) \in [0,T) \times \mathbb{R}$,

$$\begin{aligned} u_t(x) = &\int_{\mathbb{R}} p_{T-t}(x-y) u_T(y)\,\mathrm{d}y - \int_t^T \int_{\mathbb{R}} p_{r-t}(x-y) f_r(y)\,\mathrm{d}y\,\mathrm{d}r \\ &+ \int_t^T \int_{\mathbb{R}} \partial_x p_{r-t}(x-y) \int_x^y \partial_x u_r(z)\,\mathrm{d}Y_r(z)\,\mathrm{d}y\,\mathrm{d}r. \end{aligned} \tag{12}$$

We emphasize that a notion of weak solution could be given as well, but we won't use it.

Remark 4. *When $(W^T, \mathscr{W}^T)$ is geometric, the last term in the right-hand side coincides (by integration by parts, which is made licit by approximation by smooth paths and by exponential growth of $\partial_x u$ and polynomial growth of the rough path norm of $(W^T = (Y, Z^T), \mathscr{W}^T)$) with*

$$\int_t^T \int_{\mathbb{R}} p_{r-t}(x-y) \partial_x u_r(y)\,\mathrm{d}Y_r(y)\,\mathrm{d}r,$$

which reads as a more 'natural formulation' of a mild solution and which is, by the way, the formulation used in Sections 3.1 and 3.2 of Hairer [14] *in the analysis of the KPZ equation. The point is that the formulation* (12) *seems a bit more tractable as it splits into two well separated parts the rough integration and the regularization effect of the heat kernel. Once again, both are equivalent in the geometric (and in particular smooth) setting.*

Here is a crucial result in our analysis (the proof is postponed to Section 3):

Theorem 5. *Suppose that Y verifies the conditions of the previous definition. Then for any $f \in L^\infty((0,T) \times \mathbb{R})$ and any $u^T \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$ such that $\sup_{r \geqslant 1} e^{-\vartheta r} [\![(u^T)']\!]_\beta^{[-r,r]} < \infty$, there exists a unique mild solution to the problem $\mathcal{P}(Y,f,T)$ with the terminal condition $u_T = u^T$. Moreover, letting $\rho = \max\big[1, T, \vartheta, \|f\|_\infty, \sup_{r \geqslant 1} e^{-\vartheta r} [\![(u^T)']\!]_\beta^{[-r,r]}, \kappa_{\alpha,\chi}(W^T, \mathscr{W}^T)\big]$, we can find a constant $C = C(\rho, \alpha, \beta, \chi)$, such that, for any $(t,x) \in [0,T] \times \mathbb{R}$,*

$$|u_t(x)| + |\partial_x u_t(x)| \leqslant C \exp\big(C|x|\big). \tag{13}$$

and for any $(s,t,x,y) \in [0,T]^2 \times \mathbb{R}^2$,

$$\begin{aligned} &|u_t(x) - u_s(x)| \leqslant C \exp\big(C|x|\big) |t-s|^{\frac{1+\beta}{2}}, \\ &|\partial_x u_t(x) - \partial_x u_s(y)| \leqslant C \exp\big(C[|x| \vee |y|]\big) \big(|t-s|^{\frac{\beta}{2}} + |x-y|^\beta\big). \end{aligned} \tag{14}$$

As already explained, it is then quite natural to wonder about the stability of mild solutions under mollification of $(W^T, \mathscr{W}^T)$. In that framework, it is worth specifying the mollification strategy. A 'physical' way for mollifying W^T consists indeed in mollifying Y in x first - the mollification is then infinitely differentiable in x, the derivatives being continuous in space and time- and then in replacing Y by its mollified version in (11). Denoting by Y^n the mollified path at the nth step of the mollified sequence, the resulting $Z^{n,T}$ is smooth in x, the derivatives being also continuous in space and time. This permits to define the corresponding pair $(W^{n,T}, \mathscr{W}^{n,T})$ directly. In that specific *geometric* setting, we claim (once again, the proof is deferred to Section 3):

Proposition 6. *In the same framework as in Theorem 5, assume that the rough path $(W^T, \mathscr{W}^T)$ is geometric in the sense that there exists a sequence of smooth paths $(Y^n)_{n\geqslant 1}$ such that the corresponding sequence $(W^{n,T} = (Y^n, Z^{n,T}))_{n\geqslant 1}$ satisfies*

(1) $\|(W^T - W^{n,T}, \mathscr{W}^T - \mathscr{W}^{n,T})\|_{0,\alpha}^{[0,T)\times\mathbb{I}}$ *tends to 0 as n tends to ∞ for any segment $\mathbb{I}\subset\mathbb{R}$, where $\mathscr{W}_t^{n,T}(x,y) = \int_x^y (W^{n,T}(z) - W^{n,T}(x)) \otimes \mathrm{d}W^{n,T}(z)$, for $t\in[0,T)$ and $x,y\in\mathbb{R}$,*

(2) $\sup_{n\geqslant 1} \kappa_{\alpha,\chi}((W_t^{n,T}, \mathscr{W}_t^{n,T})_{0\leqslant t\leqslant T})$ *is finite (see (9) for the definition of κ_χ).*

Then, the associated solutions $(u^n)_{n\geqslant 1}$ (in the sense of Definition 3) and their gradients in space $(v^n = \partial_x u^n)_{n\geqslant 1}$ converge towards u and $v = \partial_x u$ uniformly on compact subsets of $[0,T]\times\mathbb{R}$.

2.6. **Martingale problem.** We can now give a rigorous definition of the martingale problem associated with (1):

Definition 7. *Let $T_0 > 0$ and $x_0\in\mathbb{R}$. Given $Y\in\mathcal{C}([0,T_0)\times\mathbb{R},\mathbb{R})$, assume that, for any $0\leqslant T\leqslant T_0$, there exists $\mathscr{W}^T$ such that $(W^T = (Y, Z^T), \mathscr{W}^T)$ belongs to $\mathcal{R}^{\alpha,\chi}([0,T)\times\mathbb{R},\mathbb{R}^2)$ with $\alpha > 1/3$ and $\chi < \alpha/2$, the supremum $\sup_{0\leqslant T\leqslant T_0}\kappa_{\alpha,\chi}((W_t^T, \mathscr{W}_t^T)_{0\leqslant t<T})$ being finite.*

A probability measure $\mathbb{P}$ on $\mathcal{C}([0,T_0],\mathbb{R})$ (endowed with the canonical filtration $(\mathcal{F}_t)_{0\leqslant t\leqslant T_0}$) is said to solve the martingale problem related to $\mathcal{L}$ starting from x if the canonical process $(X_t)_{0\leqslant t\leqslant T_0}$ satisfies the following two conditions:

(1) $\mathbb{P}(X_0 = x_0) = 1$,

(2) *for any $T\in[0,T_0]$ and any locally Hölder continuous and bounded function $f:(0,T)\times\mathbb{R}\to\mathbb{R}$, the process $(u_t(X_t) - \int_0^t f_r(X_r)\,\mathrm{d}r)_{0\leqslant t\leqslant T}$ is a square integrable martingale under $\mathbb{P}$, where u is a mild solution of $\mathcal{P}(Y,f,T)$ (with a given value for u_T).*

A similar definition holds by letting the canonical process start from x_0 at some time $t_0\neq 0$, in which case we say that the initial condition is (t_0,x_0) and (1) is replaced by $\mathbb{P}(\forall s\in[0,t_0],\ X_s = x_0) = 1$.

Pay attention that we require more in Definition 7 than in Definition 3 as we let the terminal time T vary within the interval $[0,T_0]$. In particular, for considering a solution to the martingale problem, it is not enough to assume that, at terminal time T_0, $(W^{T_0}, \mathscr{W}^{T_0})$ belongs to $\mathcal{R}^{\alpha,\chi}([0,T_0)\times\mathbb{R},\mathbb{R}^2)$. The rough path structure must also exist at any $0\leqslant T<T_0$, the regularity of the path W^T and of its iterated integral $\mathcal{W}^T$ being uniformly controlled in $T\in[0,T_0]$.

Our goal is then to prove existence and uniqueness of a solution:

Theorem 8. *In addition to the assumption of Definition 7, assume that, at any time $0\leqslant T\leqslant T_0$, $(W^T, \mathscr{W}^T)$ is geometric (in the sense of Proposition 6), the paths $(Y^n)_{n\geqslant 1}$ used for defining the approximating paths $(W^{n,T}, \mathscr{W}^{n,T})_{n\geqslant 1}$ being the same for all the T's and the supremum $\sup_{0\leqslant T\leqslant T_0}\sup_{n\geqslant 1}\kappa_{\alpha,\chi}((W_t^{n,T}, \mathscr{W}_t^{n,T})_{0\leqslant t<T})$ being finite. Then, given an initial condition x_0, the martingale problem has a unique solution.*

Remark 9. *The martingale problem is here set on the finite interval $[0,T_0]$. Obviously, existence and uniqueness extend to $[0,\infty)$. Notice also that β doesn't play any role in the existence and uniqueness of a solution. It will play a role when discussing the dynamics of the solution in Section 4.*

2.7. **Proof of the solvability of the martingale problem.** *First step.* We consider a sequence of paths $(Y^n)_{n\geqslant 1}$ as in the statement of Proposition 6. Replacing $(Y^n_t(x))_{0\leqslant t\leqslant T, x\in\mathbb{R}}$ by $(\varphi^n(Y^n_t(x)))_{0\leqslant t\leqslant T, x\in\mathbb{R}}$ (and modifying accordingly the definition of $(W^{n,T}, \mathscr{W}^{n,T})_{n\geqslant 1}$, for $0 \leqslant T \leqslant T_0$), for some suitable compactly supported smooth mapping φ^n that converges toward the identity on compact subsets of $\mathbb{R}$ as n tends to ∞, we can assume (without any loss of generality) that Y^n has bounded derivatives on the whole space.

We then notice that, for a given $x_0 \in \mathbb{R}$, the SDE (set on some filtered probability space endowed with a Brownian motion $(B_t)_{0\leqslant t\leqslant T_0}$)

$$\mathrm{d}X^n_t = \mathrm{d}B_t + \partial_x Y^n_t(X^n_t)\,\mathrm{d}t, \quad t\in[0,T_0] \quad ; \quad X_0 = x_0, \tag{15}$$

admits a unique solution.

Second step. Choosing $\beta \in (1/3, \alpha)$ with $\beta > 2\chi$ and letting $u^T(x) = \exp(\vartheta x)$ for a given $T \in [0, T_0]$, we denote by $(u^n_t(x))_{0\leqslant t\leqslant T, x\in\mathbb{R}}$ the mild solution to (12) with $f = 0$ and Y replaced by Y^n. It is well-known that u^n is a classical solution of

$$\partial_t u^n_t(x) + \tfrac{1}{2}\partial^2_{xx} u^n_t(x) + \partial_x Y^n_t(x)\partial_x u^n_t(x) = 0, \tag{16}$$

(Eq. (12), with u replaced by u^n, can be differentiated twice in x by transferring one derivative from the heat kernel onto the integral driven by $\mathrm{d}Y^n$; then, the second order derivative can be proved to be continuous in space and time), so that, by Itô's formula, the process $(u^n_t(X^n_t))_{0\leqslant t\leqslant T}$ is a true martingale (since we know, from Theorem 5, that u^n is at most of exponential growth). Then, (13) yields

$$\mathbb{E}\big[\exp\big(\vartheta X^n_T\big)\big] = \mathbb{E}\big[u^n_T\big(X^n_T\big)\big] = u_0(x_0) \leqslant C\exp(C|x_0|),$$

where $C = C(\alpha, \beta, \chi, \rho)$ as in Theorem 5. A crucial thing is that ρ is uniformly bounded in $T \in [0, T_0]$ so that it can be assumed to be independent of T. Replacing $u^T(x)$ by $u^T(-x)$, we get the same result with ϑ replaced by $-\vartheta$ in the above inequality, so that

$$\mathbb{E}\big[\exp\big(\vartheta|X^n_T|\big)\big] \leqslant C\exp\big(C|x_0|\big).$$

Therefore, the exponential moments of X^n_T are bounded, uniformly in $n \geqslant 1$. As C is independent of $T \in [0, T_0]$, we deduce that the marginal exponential moments of $(X^n_t)_{0\leqslant t\leqslant T_0}$ are bounded, uniformly in $n \geqslant 1$.

Third step. Now we change the domain of definition and the terminal condition of the PDE. We consider the PDE on $[0, t+h]\times\mathbb{R}$ with $u^{t+h}(x) = x$ as boundary condition, where $0 \leqslant t \leqslant t+h \leqslant T_0$. To simplify, we still denote by $(u^n_s(x))_{0\leqslant s\leqslant t+h, x\in\mathbb{R}}$ the mild solution to (12) with $f = 0$, Y replaced by Y^n and $u^n_{t+h} = u^{t+h}$ as terminal condition. By Itô's formula,

$$\begin{aligned} X^n_{t+h} - X^n_t &= u^n_{t+h}(X^n_{t+h}) - u^n_t(X^n_t) + u^n_t(X^n_t) - u^n_{t+h}(X^n_t) \\ &= \int_t^{t+h} \partial_x u^n_s(X^n_s)\,\mathrm{d}B_s + u^n_t(X^n_t) - u^n_{t+h}(X^n_t). \end{aligned} \tag{17}$$

Therefore, by (13) and (14), we deduce that, for any $q \geqslant 1$, there exists a constant C_q, independent of n, such that

$$\begin{aligned}\mathbb{E}\big[|X^n_{t+h} - X^n_t|^q\big]^{\frac{1}{q}} &\leqslant C_q\Big\{\mathbb{E}\Big[\Big(\int_t^{t+h} |\partial_x u^n_s(X^n_s)|^2\,\mathrm{d}s\Big)^{\frac{q}{2}}\Big]^{\frac{1}{q}} + \mathbb{E}\big[|u^n_t(X^n_t) - u^n_{t+h}(X^n_t)|^q\big]^{\frac{1}{q}}\Big\}\\ &\leqslant C_q\big\{h^{\frac{1}{2}-\frac{1}{q}} \sup_{0\leqslant s\leqslant T_0} \mathbb{E}\big[|\partial_x u^n_s(X^n_s)|^q\big]^{\frac{1}{q}} + \mathbb{E}\big[|u^n_t(X^n_t) - u^n_{t+h}(X^n_t)|^q\big]^{\frac{1}{q}}\big\}\\ &\leqslant C_q\big\{h^{\frac{1}{2}-\frac{1}{q}} \sup_{0\leqslant s\leqslant T_0} \mathbb{E}\big[\exp(q|X^n_s|)\big]^{\frac{1}{q}} + h^{\frac{1+\beta}{2}} \sup_{0\leqslant s\leqslant T_0} \mathbb{E}\big[\exp(q|X^n_s|)|\big]^{\frac{1}{q}}\big\}.\end{aligned}$$

By the second step (uniform boundedness of the exponential moments) and by Kolmogorov's criterion, we deduce that the processes $(X^n_t)_{0\leqslant t\leqslant T_0}$ are tight.

Fourth step. It remains to prove that any weak limit $(X_t)_{0\leqslant t\leqslant T_0}$ is a solution to the martingale problem. For a given $T \in [0, T_0]$, we know from Proposition 6 that we can find a sequence $(u^n)_{n\geqslant 1}$ of classical solutions to the problems $\mathcal{P}(Y^n, f, T)$ such that the sequence $(u^n, \partial_x u^n)_{n\geqslant 1}$ converges towards $(u, \partial_x u)$, uniformly on compact subsets of $[0, T] \times \mathbb{R}$. Note that the solutions are 'classical' as f is locally Hölder continuous (the argument is the same as for (16): Eq. (12), with u replaced by u^n, can be differentiated twice in x). Applying Itô's formula to each $(u^n_t(X^n_t))_{0\leqslant t\leqslant T}$, $n \geqslant 1$, we deduce that

$$u^n_t(X^n_t) - u^n_0(X^n_0) - \int_0^t f_s(X^n_s)\,\mathrm{d}s = \int_0^t \partial_x u^n_s(X^n_s)\,\mathrm{d}B_s, \quad 0 \leqslant t \leqslant T.$$

By (13), we know that the functions $(\partial_x u^n)_{n\geqslant 1}$ are at most of exponential growth, uniformly in $n \geqslant 1$. Moreover, we recall that the processes $((X^n_t)_{0\leqslant t\leqslant T})_{n\geqslant 1}$ have finite marginal exponential moments, uniformly in $n \geqslant 1$ as well. Therefore, the martingales $((u^n_t(X^n_t) - u^n_0(X^n_0) - \int_0^t f_s(X^n_s)\,\mathrm{d}s)_{0\leqslant t\leqslant T})_{n\geqslant 1}$ are bounded in L^2, uniformly in $n \geqslant 1$. Letting n tend to the infinity, we complete the proof.

2.8. **Well-posedness of the martingale problem.** Here is the uniqueness part in Theorem 8:

Theorem 10. *Given $T_0 > 0$, assume that the assumption of Theorem 8 is in force. For an initial condition $(t_0, x_0) \in [0, T_0] \times \mathbb{R}$, there exists a unique solution to the martingale problem (on $[0, T_0]$) with (t_0, x_0) as initial condition. It is denoted by $\mathbb{P}_{t_0,x_0}$. The mapping $[0, T_0] \times \mathbb{R} \ni (t, x) \mapsto \mathbb{P}_{t,x}(A)$ is measurable for any Borel subset A of the canonical space $\mathcal{C}([0, T_0], \mathbb{R})$. Moreover, it is strong Markov.*

Proof. Existence has been already proven in Theorem 8.

First Step. We first establish uniqueness of the marginal laws. Assume indeed that $\mathbb{P}_1$ and $\mathbb{P}_2$ are two solutions of the martingale problem with the same initial condition (t_0, x_0). Then, for any bounded and locally Hölder continuous function $f : [0, T_0) \times \mathbb{R} \to \mathbb{R}$, it holds

$$\forall t \geqslant t_0, \quad \mathbb{E}_1 \int_{t_0}^{T_0} f_s(X_s)\,\mathrm{d}s = \mathbb{E}_2 \int_{t_0}^{T_0} f_s(X_s)\,\mathrm{d}s, \tag{18}$$

where $\mathbb{E}_1$ and $\mathbb{E}_2$ denote the expectations under $\mathbb{P}_1$ and $\mathbb{P}_2$ ($(X_t)_{0\leqslant t\leqslant T_0}$ denotes the canonical process). Indeed, denoting by u the solution of the PDE $\mathcal{P}(Y, f, T_0)$ with 0 as terminal condition at time T_0, we know from the definition of the martingale problem that, both

under $\mathbb{P}_1$ and $\mathbb{P}_2$, the process $(u_s(X_s) - \int_{t_0}^s f_r(X_r)\,\mathrm{d}r)_{t_0 \leqslant s \leqslant T_0}$ is a martingale. Therefore, taking the expectation under $\mathbb{E}_1$ and $\mathbb{E}_2$ and noticing that $u_{T_0}(X_{T_0}) = 0$ almost surely under $\mathbb{P}_1$ and $\mathbb{P}_2$, we deduce that both sides in (18) are equal to $-u_{t_0}(x_0)$, which is enough to complete the proof of (18) and thus to prove that the marginal laws of the canonical process are the same under $\mathbb{P}_1$ and $\mathbb{P}_2$.

By Theorems 4.2 and 4.6 in [7], we deduce that the martingale problem has a unique solution (note that the results in [7] hold for time homogeneous martingale problems whereas the martingale problem we are here investigating is time inhomogeneous; adding an additional variable in the state space, the problem we are considering can be easily turned into a time-homogeneous one). Measurability and strong Markov property are proved as in [7]. □

3. Solving the PDE

This section is devoted to the proof of Theorem 5. As the definition of a mild solution in Definition 3 consists in a convolution of a rough integral with the heat kernel, the first step is to investigate the smoothing effect of a Gaussian kernel onto a rough integral. Existence and uniqueness of a mild solution to (12) is then proved by means of a contraction argument.

Parts of the results presented here are variations of the ones obtained in Sections 3.1 and 3.2 of Hairer [14] for solving the KPZ equation, but differ slightly in the very construction of a mild solution, see Remark 4.

3.1. Mild solutions as Picard's fixed points.

In this subsection, we fix $\alpha, \beta, \chi, \vartheta, \lambda$ such that $1/3 < \beta < \alpha \leqslant 1$, $\chi < \beta/2$ and $\vartheta, \lambda \geqslant 1$. Given $Y \in \mathcal{C}([0,T) \times \mathbb{R}, \mathbb{R})$ for some final time $T \leqslant 1$, we assume that there exists $\mathscr{W}^T$ such that $(W_t^T = (Y_t, Z_t^T), \mathscr{W}_t^T)_{0 \leqslant t \leqslant T}$ is in $\mathcal{R}^{\alpha,\chi}([0,T) \times \mathbb{R}, \mathbb{R}^2)$, $(Z_t^T)_{0 \leqslant t \leqslant T}$ being given by (11). We will simply denote by κ to the semi norm $\kappa_{\alpha,\chi}((W_t^T, \mathscr{W}_t^T)_{t \in [0,T)})$ and we will omit the superscript T in Z^T, W^T and $\mathscr{W}^T$. We also recall the definition of $\Theta_T^{\vartheta,\lambda}(v)$ for $v \in \mathcal{B}^{\beta,\vartheta}([0,T) \times \mathbb{R}, W)$:

$$\Theta_T^{\vartheta,\lambda}(v) = \sup_{\substack{r \geqslant 1 \\ t \in [0,T)}} \Big[\frac{1}{E_T^{\vartheta,\lambda}(t,r)} \Big([\![v]\!]_{\beta/2,\beta}^{[t,T) \times [-r,r]} + \tfrac{1}{2} [\![\partial_W v]\!]_{\beta/2,\beta}^{[t,T) \times [-r,r]} + \big((T-t)^{-\frac{\beta}{2}} + r^\beta \big) \|\mathscr{R}^{v_t}\|_{2\beta}^{[-r,r]} \Big) \Big],$$

with $E_T^{\vartheta,\lambda}(t,r) = \exp[\lambda(T-t) + \vartheta r(1+T-t)]$. We start with the following technical lemma, which plays a crucial role in the proof of Theorem 5:

Lemma 11. *For any $\gamma_1 \leqslant \gamma_2 \leqslant \beta/2$ and $k \in \mathbb{N}^*$, there is a constant $C = C(\alpha, \beta, \gamma_1, \gamma_2, \chi, k)$ (independent of ϑ and λ) such that for any $t, \tau \in [0,T)$, with $\tau \leqslant T - t$, and any $r \geqslant 1$, the following bounds hold for any $v \in \mathcal{B}^{\beta,\vartheta}([0,T) \times \mathbb{R}, W)$ and any $x \in [-r,r]$:*

$$\int_{\mathbb{R}} \int_0^\tau \frac{|\partial_x^k p_1(y)|}{s^{1+\gamma_1}} \left| \int_x^{x-\sqrt{s}y} v_{t+s}(z)\,\mathrm{d}Y_{t+s}(z) \right| \mathrm{d}s\,\mathrm{d}y \leqslant \Psi \lambda^{\frac{\beta-\alpha}{2}} \tau^{\gamma_2-\gamma_1} r^{\gamma_2},$$

$$\int_{\mathbb{R}} \int_0^\tau \frac{|\partial_x^k p_1(y)|}{s^{1+2\gamma_1}} \left| \int_x^{x-\sqrt{s}y} \big(v_{t+s}(z) - v_{t+s}(x) \big)\,\mathrm{d}Y_{t+s}(z) \right| \mathrm{d}s\,\mathrm{d}y \leqslant \Psi \lambda^{\frac{\beta-\alpha}{4}} \tau^{\beta-2\gamma_1} \left(r^\beta + (T-t)^{-\frac{\beta}{2}} \right),$$

with $\Psi = C e^{CT\vartheta^2} \kappa \Theta_T^{\vartheta,\lambda}(v) E_T^{\vartheta,\lambda}(t,r)$.

Proof. In the whole proof, we just denote $\Theta_T^{\vartheta,\lambda}(v)$ and $E_T^{\vartheta,\lambda}(t,r)$ by Θ and $E(t,r)$. We start with the proof of the first inequality. The point is to apply the second inequality in Lemma 2 with y replaced by $x-\sqrt{s}y$ and thus r replaced by $r+|y|$. We get

$$\left|\int_x^{x-\sqrt{s}y} v_{t+s}(z)\,\mathrm{d}Y_{t+s}(z)\right| \leqslant C\kappa\Theta E(t+s,r+|y|)\big[s^{\frac{\alpha}{2}}|y|^\alpha\big(r+|y|\big)^\chi + \mathscr{D}\big(t+s,r+|y|,\sqrt{s}y\big)\big],$$

where $C=C(\alpha,\beta)$. Noting that $E(t+s,r+|y|)\leqslant \exp[-(\lambda+\vartheta(r+|y|))s+\vartheta(1+T)|y|)]E(t,r)$ and that $\mathscr{D}(t+s,r+|y|,\sqrt{s}y)\leqslant C(1+|y|^3)\mathscr{D}(t+s,r+|y|,\sqrt{s})$, we deduce that

$$\begin{aligned}(19)\qquad &(r+|y|)^{-\gamma_2}\int_0^\tau s^{-1-\gamma_1}\left|\int_{x'}^{x'-\sqrt{s}y} v_{t+s}(z)\,\mathrm{d}Y_{t+s}(z)\right|\mathrm{d}s\\ &\qquad\leqslant C\kappa\Theta E(t,r)e^{\vartheta(1+T)|y|}\big(1+|y|^3\big)\int_0^\tau \frac{e^{-(\lambda+\vartheta(r+|y|))s}}{s^{\gamma_1}(r+|y|)^{\gamma_2}}\mathscr{D}'\big(t,s,r+|y|\big)\,\mathrm{d}s,\end{aligned}$$

where

$$(20)\qquad \mathscr{D}'(t,s,\rho)=s^{\frac{\alpha}{2}-1}\rho^\chi+s^{\alpha-1}\rho^{2\chi}+s^{\alpha+\frac{\beta}{2}-1}\rho^{2\chi+\frac{\beta}{2}}+s^{\frac{\alpha}{2}+\beta-1}\rho^\chi\left(\rho^\beta+(T-t-s)^{-\frac{\beta}{2}}\right).$$

We thus have to bound integrals of the form $\rho^{b-\gamma_2}\int_0^\tau e^{-(\lambda+\vartheta\rho)s}s^{a-\gamma_1-1}\,\mathrm{d}s$ with $a\geqslant\alpha/2$ ($\geqslant\gamma_2$), $0<b\leqslant a$ and $\rho\geqslant 1$. Bounding $s^{\gamma_2-\gamma_1}$ by $\tau^{\gamma_2-\gamma_1}$ and noticing that

$$\begin{aligned}(21)\qquad \frac{\rho^{b-\gamma_2}}{(\lambda+\vartheta\rho)^{a-\gamma_2}}&\leqslant\frac{\rho^{b-\gamma_2}}{(\lambda+\rho)^{a-\gamma_2}}\\ &\leqslant \rho^{b-a}\mathbf{1}_{\{\rho\geqslant\lambda\}}+\lambda^{\gamma_2-a}\mathbf{1}_{\{1\leqslant\rho<\lambda\}}\mathbf{1}_{\{b<\gamma_2\}}+\lambda^{b-a}\mathbf{1}_{\{\rho<\lambda\}}\mathbf{1}_{\{b\geqslant\gamma_2\}}\leqslant\lambda^{(b\vee\gamma_2)-a},\end{aligned}$$

we get the following upper bound for the integral (performing a change of variable to pass from the first to the second line and recalling that $\gamma_2\leqslant\beta/2$ to derive the last inequality):

$$\begin{aligned}(22)\qquad \rho^{b-\gamma_2}\int_0^\tau e^{-(\lambda+\vartheta\rho)s}s^{a-\gamma_1-1}\,\mathrm{d}s&\leqslant\tau^{\gamma_2-\gamma_1}\rho^{b-\gamma_2}\int_0^\tau e^{-(\lambda+\vartheta\rho)s}s^{a-\gamma_2-1}\,\mathrm{d}s\\ &\leqslant\frac{\tau^{\gamma_2-\gamma_1}\rho^{b-\gamma_2}}{(\lambda+\vartheta\rho)^{a-\gamma_2}}\int_0^{(\lambda+\vartheta\rho)\tau}e^{-s}s^{a-\gamma_2-1}\,\mathrm{d}s\leqslant\tau^{\gamma_2-\gamma_1}\lambda^{(b\vee\frac{\beta}{2})-a}\Gamma(a-\gamma_2).\end{aligned}$$

Because of the term in $(T-t-s)$ in the definition of $\mathscr{D}'$, we also have to control

$$\begin{aligned}(23)\qquad &\rho^{\chi-\gamma_2}\int_0^\tau\frac{e^{-(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}-\beta+\gamma_1}(T-t-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\leqslant\tau^{\gamma_2-\gamma_1}\rho^{\chi-\gamma_2}\int_0^\tau\frac{e^{-(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}-\beta+\gamma_2}(T-t-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\\ &=\tau^{\gamma_2-\gamma_1}\frac{\rho^{\chi-\gamma_2}}{(\lambda+\vartheta\rho)^{\frac{\alpha}{2}-\gamma_2}}\frac{\tau^{\frac{\beta}{2}}}{(T-t)^{\frac{\beta}{2}}(\lambda+\vartheta\rho)^{\frac{\beta}{2}}}\int_0^1\frac{(\tau(\lambda+\vartheta\rho))^{\frac{\alpha}{2}+\frac{\beta}{2}-\gamma_2}e^{-\tau(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}-\beta+\gamma_2}[1-\tau s/(T-t)]^{\frac{\beta}{2}}}\,\mathrm{d}s.\end{aligned}$$

In order to bound the integral in the second line, we make use of the inequality $x^ae^{-xs}\leqslant a^ae^{-a}/s^a$, which holds for $s\in(0,1]$ and $a,x\geqslant 0$. Using also the bounds $\tau\leqslant T-t$ and $\lambda+\vartheta\rho\geqslant 1$ together with (21), we get (for a possibly new value of the constant C):

$$\begin{aligned}(24)\qquad &\rho^{\chi-\gamma_2}\int_0^\tau\frac{e^{-(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}-\beta+\gamma_1}(T-t-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\\ &\qquad\leqslant C\tau^{\gamma_2-\gamma_1}\lambda^{(\chi\vee\gamma_2)-\frac{\alpha}{2}}\int_0^1\frac{\mathrm{d}s}{s^{1-\frac{\beta}{2}}(1-s)^{\frac{\beta}{2}}}\leqslant C\tau^{\gamma_2-\gamma_1}\lambda^{\frac{\beta-\alpha}{2}}.\end{aligned}$$

A careful inspection of (20) shows that we can apply (22) and (24) with $a \geqslant \alpha/2$ and $b-a \leqslant \chi-\alpha/2$ in order to bound (19) (a is the part different from -1 in the exponent of s and b is the exponent of ρ). We obtain

$$(r+|y|)^{-\gamma_2}\int_0^\tau s^{-1-\gamma_1}\left|\int_x^{x-\sqrt{s}y} v_{t+s}(z)\,\mathrm{d}Y_{t+s}(z)\right|\mathrm{d}s \leqslant C\kappa\Theta E(t,r)e^{\vartheta(1+T)|y|}\big(1+|y|^3\big)\tau^{\gamma_2-\gamma_1}\lambda^{\frac{\beta-\alpha}{2}}. \tag{25}$$

As $r^{-\gamma_2}\leqslant(1+|y|)^{\gamma_2}(r+|y|)^{-\gamma_2}$, we get the first bound of the lemma by integrating (25) against $\big|\partial_x^k p_1(y)\big|$.

We now turn to the proof of the second inequality in the statement. We make use of the first inequality in Lemma 2. Replacing $v_{t+s}(z)$ by $v_{t+s}(z)-v_{t+s}(x)$ in (19), we get the same inequality but with a simpler form of $\mathcal{D}'(t,s,r+|y|)$, namely the first term in the right-hand side in (20) doesn't appear. This says that we can now apply (22) with $a\geqslant\alpha\wedge(\alpha/2+\beta)\geqslant\beta$ and $b-a\leqslant\chi-\alpha/2$. The value of a being larger than β, this permits to apply (21) with γ_2 replaced by $2\gamma_2$. Then, we can replace γ_1 and γ_2 by $2\gamma_1$ and $2\gamma_2$ in (22). With the prescribed values of a and b, the resulting bound in (22) is $C\tau^{2\gamma_2-2\gamma_1}\lambda^{(b\vee\beta)-a}$. Choosing $\gamma_2=\beta/2$ and following (25), we see that the contribution of (22) in the second inequality of the statement is $\Psi\lambda^{(\beta-\alpha)/2}\tau^{\beta-2\gamma_1}r^\beta\leqslant\Psi\lambda^{(\beta-\alpha)/4}\tau^{\beta-2\gamma_1}r^\beta$, which fits the first part of the inequality. To recover the second part of the inequality, we must discuss the contribution of (23). Going back to (20), we are to analyze (pay attention that, in comparison with (23), γ_2 is set to 0):

$$\begin{aligned}
&(T-t)^{\frac{\beta}{2}}\rho^\chi\int_0^\tau\frac{e^{-(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}-\beta+2\gamma_1}(T-t-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\\
&\leqslant\tau^{\beta-2\gamma_1}\rho^\chi\int_0^\tau\frac{e^{-(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}}(1-s/(T-t))^{\frac{\beta}{2}}}\,\mathrm{d}s\leqslant\tau^{\beta-2\gamma_1}\tau^{\alpha/2}\rho^\chi\int_0^1\frac{e^{-\tau(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}}(1-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\\
&=\tau^{\beta-2\gamma_1}\tau^{\frac{\alpha/2-\chi}{2}}\frac{\rho^\chi}{(\lambda+\vartheta\rho)^{\frac{\alpha/2+\chi}{2}}}\int_0^1\frac{(\tau(\lambda+\vartheta\rho))^{\frac{\alpha/2+\chi}{2}}e^{-\tau(\lambda+\vartheta\rho)s}}{s^{1-\frac{\alpha}{2}}(1-s)^{\frac{\beta}{2}}}\,\mathrm{d}s\leqslant C\lambda^{\frac{\chi-\alpha/2}{2}}\tau^{\beta-2\gamma_1},
\end{aligned} \tag{26}$$

the last inequality following from (21). Noticing that $\chi<\beta/2$, this gives the second part of the second inequality of the statement. □

Here is now the key result to prove Theorem 5.

Theorem 12. *Keep the notations and assumptions introduced at the beginning of Subsection 3.1. For $(v,\partial_W v)\in\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$, define the function $\mathcal{M}(v,\partial_W v):[0,T)\times\mathbb{R}\to\mathbb{R}$ together with its W-derivative by letting, for any $t\in[0,T)$ and $x\in\mathbb{R}$,*

$$\big[\mathcal{M}(v,\partial_W v)\big]_t(x)=\int_t^T\int_{\mathbb{R}}\partial_x^2 p_{s-t}(x-y)\int_x^y v_s(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s.$$

$$\partial_W\big[\mathcal{M}(v,\partial_W v)\big]_t(x)=\big(0,v_t(x)\big)\quad(\textit{i.e. }\partial_Y\mathcal{M}(v,\partial_W v)_t(x)=0,\ \partial_Z\mathcal{M}(v,\partial_W v)_t(x)=v_t(x)).$$

(With an abuse of notation, we will just write $(\mathcal{M}v)_t(x)$ for $[\mathcal{M}(v,\partial_W v)]_t(x)$.) Then $\mathcal{M}$ defines a bounded operator from $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$ into itself. Moreover, there exists a positive constant $C=C(\alpha,\beta,\chi)$ such that for every $v\in\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$,

$$\Theta_T^{\vartheta,\lambda}(\mathcal{M}v)\leqslant\big(\tfrac12+C\kappa\exp(CT\vartheta^2)\lambda^{-\epsilon}\big)\Theta_T^{\vartheta,\lambda}(v),\qquad\textit{with }\epsilon:=(\alpha-\beta)/4.$$

Proof. As in the proof of Lemma 11, we just denote $\Theta_T^{\vartheta,\lambda}(v)$ and $E_T^{\vartheta,\lambda}(t,r)$ by Θ and $E(t,r)$. By an obvious change of variable, we get for any $r \geqslant 1$, $x \in [-r,r]$ and $t \in [0,T)$,

$$(\mathcal{M}v)_t(x) = \int_{\mathbb{R}} \partial_x^2 p_1(y) \int_0^{T-t} s^{-1} \int_x^{x-\sqrt{s}y} v_{t+s}(z)\, \mathrm{d}Y_{t+s}(z)\, \mathrm{d}s\, \mathrm{d}y.$$

Then the first inequality of Lemma 11 with $\gamma_1 = \gamma_2 = 0$, $\tau = T-t$ and $k = 2$ leads to

$$\big(E(t,r)\big)^{-1} \left|(\mathcal{M}v)_t(x)\right| \leqslant C\kappa e^{CT\vartheta^2} \lambda^{\frac{\beta-\alpha}{2}} \Theta, \tag{27}$$

where $C = C(\alpha,\beta,\chi)$.

We now study the time variations of $\mathcal{M}v$. For $0 \leqslant t \leqslant s \leqslant T$ and $x \in \mathbb{R}$, we deduce from the identity $\frac{1}{2}\partial_x^2 p = \partial_t p$:

$$\begin{aligned}
\left|(\mathcal{M}v)_s(x) - (\mathcal{M}v)_t(x)\right| &\leqslant \frac{1}{2}\left|\int_s^T \int_t^s \int_{\mathbb{R}} \partial_x^4 p_{\rho-u}(x-y) \int_x^y v_\rho(z)\, \mathrm{d}Y_\rho(z)\, \mathrm{d}y\, \mathrm{d}u\, \mathrm{d}\rho\right| \\
&\quad + \left|\int_t^s \int_{\mathbb{R}} \partial_x^2 p_{\rho-t}(x-y) \int_x^y v_\rho(z)\, \mathrm{d}Y_\rho(z)\, \mathrm{d}y\, \mathrm{d}\rho\right| \\
&:= \frac{1}{2}\mathcal{T}_1 + \mathcal{T}_2.
\end{aligned}$$

By the changes of variable $(\rho,u) \mapsto (s+\rho-u, s-u)$ and then $y \mapsto x - \sqrt{\rho}s$, we get:

$$\begin{aligned}
\mathcal{T}_1 &= \left|\int_{\mathbb{R}} \partial_x^4 p_1(y) \int_0^{s-t} \int_u^{T-s+u} \frac{1}{\rho^2} \int_x^{x-\sqrt{\rho}y} v_{s+\rho-u}(z)\, \mathrm{d}Y_{s+\rho-u}(z)\, \mathrm{d}\rho\, \mathrm{d}u\, \mathrm{d}y\right| \\
&\leqslant \int_{\mathbb{R}} \left|\partial_x^4 p_1(y)\right| \int_0^{s-t} u^{\frac{\beta}{2}-1} \int_0^{T-t} \frac{1}{\rho^{1+\frac{\beta}{2}}} \left|\int_x^{x-\sqrt{\rho}y} v_{s+\rho-u}(z)\, \mathrm{d}Y_{s+\rho-u}(z)\right| \mathrm{d}\rho\, \mathrm{d}u\, \mathrm{d}y.
\end{aligned}$$

Applying Lemma 11 with $\tau = T-t$, $\gamma_1 = \gamma_2 = \beta/2$ and $k = 4$, we obtain

$$r^{-\frac{\beta}{2}} \mathcal{T}_1 \leqslant Ce^{CT\vartheta^2} \kappa\Theta E(t,r) \lambda^{\frac{\beta-\alpha}{2}} \int_0^{s-t} u^{\frac{\beta}{2}-1}\, \mathrm{d}u \leqslant Ce^{CT\vartheta^2} \kappa\Theta E(t,r) \lambda^{\frac{\beta-\alpha}{2}} (s-t)^{\frac{\beta}{2}},$$

where $C = C(\alpha,\beta,\chi)$. In order to handle $\mathcal{T}_2$, we can directly use Lemma 11 with $\tau = s-t$, $\gamma_1 = 0$, $\gamma_2 = \beta/2$ and $k = 2$. We then obtain the same bound as for $\mathcal{T}_1$, so that

$$r^{-\frac{\beta}{2}} \big(E(t,r)\big)^{-1} \left|(\mathcal{M}v)_s(x) - (\mathcal{M}v)_t(x)\right| \leqslant Ce^{CT\vartheta^2} \kappa\Theta \lambda^{\frac{\beta-\alpha}{2}} (s-t)^{\frac{\beta}{2}}. \tag{28}$$

We now investigate the space variations. Fix $-r \leqslant x < x' \leqslant r$. If $|x'-x|^2 \leqslant T-t$, the space increment between x and x' reads:

$$\begin{aligned}
&\left|(\mathcal{M}v)_t(x') - (\mathcal{M}v)_t(x)\right| \\
&= \left|\int_t^T \int_{\mathbb{R}} \big(\partial_x^2 p_{s-t}(x'-y) - \partial_x^2 p_{s-t}(x-y)\big) \int_x^y v_s(z)\, \mathrm{d}Y_s(z)\, \mathrm{d}y\, \mathrm{d}s\right| \\
&\leqslant \mathcal{I}_1^{x,x'}(x) + \mathcal{I}_1^{x,x'}(x') + \mathcal{I}_2^{x,x'},
\end{aligned} \tag{29}$$

with (using the fact that the mapping $\mathbb{R} \ni z \mapsto \partial_x^2 p_s(z)$ is centered)

$$\mathcal{I}_1^{x,x'}(\xi) := \left| \int_0^{|x'-x|^2} \int_{\mathbb{R}} \partial_x^2 p_s(\xi - y) \int_\xi^y v_{t+s}(z)\, \mathrm{d}Y_{t+s}(z)\, \mathrm{d}y\, \mathrm{d}s \right|,$$
$$\mathcal{I}_2^{x,x'} := \left| \int_{|x'-x|^2}^{T-t} \int_{\mathbb{R}} \int_x^{x'} \partial_x^3 p_s(u - y) \int_x^y v_{t+s}(z)\, \mathrm{d}Y_{t+s}(z)\, \mathrm{d}u\, \mathrm{d}y\, \mathrm{d}s \right|.$$

By Lemma 11 with $\tau = |x'-x|^2$, $\gamma_1 = 0$, $\gamma_2 = \beta/2$ and $k = 2$, we get

$$r^{-\frac{\beta}{2}} \big(E(t,r)\big)^{-1} \big(\mathcal{I}_1^{x,x'}(x) + \mathcal{I}_2^{x,x'}(x')\big) \leqslant C e^{CT\vartheta^2} \kappa \Theta \lambda^{\frac{\beta-\alpha}{2}} |x'-x|^\beta. \tag{30}$$

The term $\mathcal{I}_2^{x,x'}$ can be bounded in the following way:

$$\begin{aligned} \mathcal{I}_2^{x,x'} &\leqslant \int_{\mathbb{R}} \big|\partial_x^3 p_1(y)\big| \int_x^{x'} \int_{|x'-x|^2}^{T-t} s^{-\frac{3}{2}} \left| \int_u^{u-\sqrt{s}y} v_{t+s}(z)\, \mathrm{d}Y_{t+s}(z) \right| \mathrm{d}s\, \mathrm{d}u\, \mathrm{d}y \\ &\leqslant |x'-x|^{\beta-1} \int_{\mathbb{R}} \big|\partial_x^3 p_1(y)\big| \int_x^{x'} \int_{|x'-x|^2}^{T-t} s^{-1-\frac{\beta}{2}} \left| \int_u^{u-\sqrt{s}y} v_{t+s}(z)\, \mathrm{d}Y_{t+s}(z) \right| \mathrm{d}s\, \mathrm{d}u\, \mathrm{d}y. \end{aligned} \tag{31}$$

Using now Lemma 11 with $\tau = T-t$, $\gamma_1 = \gamma_2 = \beta/2$ and $k = 3$ we obtain:

$$r^{-\frac{\beta}{2}} \big(E(t,r)\big)^{-1} \mathcal{I}_2^{x,x'} \leqslant C e^{CT\vartheta^2} \kappa \Theta \lambda^{\frac{\beta-\alpha}{2}} |x'-x|^\beta.$$

We end up with the following bound for the space increment:

$$r^{-\frac{\beta}{2}} \big(E(t,r)\big)^{-1} \big|(\mathcal{M}v)_t(x') - (\mathcal{M}v)_t(x)\big| \leqslant C e^{CT\vartheta^2} \kappa \Theta \lambda^{\frac{\beta-\alpha}{2}} |x'-x|^\beta. \tag{32}$$

Recall that (32) holds true when $|x'-x|^2 \leqslant T-t$. When $|x'-x|^2 > T-t$, the argument is straightforward as the space increment is smaller $\mathcal{I}_1^{x,x'}(x)$ and $\mathcal{I}_1^{x,x}(x')$, so that (32) holds as well.

We study $\mathscr{R}^{(\mathcal{M}v)_t}$ in a similar way. Recalling the definition (7) for $\mathscr{R}^{(\mathcal{M}v)_t}$, we then make use of the very definition of Z^T, see (11):

$$\begin{aligned} \mathscr{R}^{(\mathcal{M}v)_t}(x,x') &= (\mathcal{M}v)_t(x') - (\mathcal{M}v)_t(x) - v_t(x)\big(Z_t^T(x') - Z_t^T(x)\big) \\ &= \int_t^T \int_{\mathbb{R}} \big(\partial_x^2 p_{s-t}(x'-y) - \partial_x^2 p_{s-t}(x-y)\big) \int_x^y (v_s(z) - v_t(x))\, \mathrm{d}Y_s(z)\, \mathrm{d}y\, \mathrm{d}s. \end{aligned}$$

The strategy consists of the same decomposition as the one used to prove (32) except that we now apply the second inequality in Lemma 11 and not the first one. This leads to new definitions of $\mathcal{I}_1^{x,x'}$ and $\mathcal{I}_2^{x,x'}$, the term $v_{t+s}(x)$ being subtracted to $v_{t+s}(z)$. If, instead of $v_{t+s}(x)$, $v_{t+s}(\xi)$ was subtracted to $v_{t+s}(z)$ in the definition of $\mathcal{I}_1^{x,x'}(\xi)$ and $v_{t+s}(u)$ was subtracted to $v_{t+s}(z)$ in the definition of $\mathcal{I}_2^{x,x'}$, Lemma 11 would give

$$\big((T-t)^{\frac{\beta}{2}} \wedge r^{-\beta}\big) \big(E(t,r)\big)^{-1} |\mathscr{R}^{(\mathcal{M}v)_t}(x,x')| \leqslant C e^{CT\vartheta^2} \kappa \Theta \lambda^{\frac{\beta-\alpha}{4}} |x'-x|^{2\beta}. \tag{33}$$

Therefore, we must discuss the fact that the centering term in both $\mathcal{I}_1^{x,x'}(x')$ and $\mathcal{I}_2^{x,x'}$ is $v_{t+s}(x)$. For handling $\mathcal{I}_1^{x,x'}(x')$, we must investigate

$$\left| \int_0^{|x'-x|^2 \wedge (T-t)} \int_{\mathbb{R}} \partial_x^2 p_s(x'-y) \int_{x'}^y \big(v_{t+s}(x') - v_{t+s}(x)\big)\, \mathrm{d}Y_{t+s}(z)\, \mathrm{d}y\, \mathrm{d}s \right|.$$

Letting $\tilde{v}_s(z) = v_s(x') - v_s(x)$, we see that $\Theta_T^{\vartheta,\lambda}(\tilde{v}) \leqslant r^{\beta/2}|x'-x|^\beta\Theta$. Then, we are led back to the original definition of $\mathcal{I}_1^{x,x'}(x')$ but with v_s replaced by the constant function $\tilde{v}_s$. Following (30), we then get that the above term is less than $C\exp(CT\vartheta^2)\kappa\Theta\lambda^{(\beta-\alpha)/2}r^\beta|x'-x|^{2\beta}$.

The strategy for handling the correction in $\mathcal{I}_2^{x,x'}$ is completely similar and leads to the same bound. We deduce that (33) holds true.

Finally, as the W-derivative of $(\mathcal{M}v)_t$ is defined as $\partial_W(\mathcal{M}v)_t = (0, v_t)$, we have

$$\frac{1}{2}\big(E(t,r)\big)^{-1}\|\partial_W(\mathcal{M}v)_t\|_{\beta/2,\beta}^{[t,T)\times[-r,r]} \leqslant \frac{1}{2}\Theta. \tag{34}$$

From (27), (28), (32), (33) and (34), we complete the proof. □

3.2. **Proof of Theorem 5.** *First step.* As in the previous subsection, we omit the superscript T in Z^T, W^T and $\mathscr{W}^T$. We then start with a technical remark. For any $T \leqslant 1$ and any $v \in \mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$, $\mathcal{M}v$ is always in $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$ by Theorem 12. Actually, this result remains true when $T \geqslant 1$. It is indeed well-checked that the bound $T \leqslant 1$ in the statement of Theorem 12 is only useful to get a resulting constant C that depends on a minimal number of parameters and that a similar result holds when $T \leqslant T_0$, for some $T_0 \geqslant 1$, provided the constant C is allowed to depend upon T_0.

Now, for a function $f \in L^\infty((0,T)\times\mathbb{R})$, a continuously differentiable function u^T from $\mathbb{R}$ into itself such that $\sup_{r\geqslant 1} e^{-\vartheta r}[\![(u^T)']\!]_\beta^{[-r,r]} < \infty$ and a function $v \in \mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$, we let for $(t,x) \in [0,T)\times\mathbb{R}$

$$\begin{aligned}\big(\widehat{\mathcal{M}}v\big)_t(x) &:= \int_{\mathbb{R}} \partial_x p_{T-t}(x-y)u^T(y)\,\mathrm{d}y - \int_t^T\int_{\mathbb{R}} \partial_x p_{s-t}(x-y)f_s(y)\,\mathrm{d}y\,\mathrm{d}s + \big(\mathcal{M}v\big)_t(x)\\ &= \int_{\mathbb{R}} p_{T-t}(y)(u^T)'(x-y)\,\mathrm{d}y - \int_t^T\int_{\mathbb{R}} \partial_x p_{s-t}(x-y)f_s(y)\,\mathrm{d}y\,\mathrm{d}s + \big(\mathcal{M}v\big)_t(x).\end{aligned} \tag{35}$$

The point is to check that $\widehat{\mathcal{M}}v$ can be lifted up into an element of $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$. By Theorem 12, the last part of the right-hand side is in $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$. Its derivative with respect to W is $\partial_W[\mathcal{M}v]$, as defined in the statement of Theorem 12. Moreover, by standard regularization properties of the heat kernel, the second term in the right-hand side is in $\mathcal{C}^{\gamma/2,\gamma}([0,T]\times\mathbb{R},\mathbb{R})$ for any $\gamma \in (0,1)$, with a finite Hölder norm on the whole $[0,T]\times\mathbb{R}$ (and not on compact subsets only). In particular, it can be lifted up into an element of $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$ with a zero derivative with respect to W. Finally, the first term is at most of exponential growth in x (with exponent ϑ). By standard regularization properties of the heat kernel, it is smooth on $[0,T)\times\mathbb{R}$. And investigating carefully the regularization effect of the heat kernel, it can be shown that

$$\sup_{0\leqslant t<T}\sup_{r\geqslant 1}\left\{(T-t)^{\frac{\beta}{2}}e^{-\vartheta r}\left\|\int_{\mathbb{R}} p_{T-t}(\cdot-y)(u^T)'(y)dy\right\|_{2\beta}^{[-r,r]}\right\} < \infty,$$

thus proving that $\sup_{0\leqslant t<T}\sup_{r\geqslant 1}\{(T-t)^{\beta/2}e^{-\vartheta r}\|\mathscr{R}\int_{\mathbb{R}} p_{T-t}(\cdot-y)(u^T)'(y)dy\|_{2\beta}^{[-r,r]}\} < \infty$, so that $\widehat{\mathcal{M}}v \in \mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R}, W)$, with $[\partial_W(\widehat{\mathcal{M}}v)]_t(x) = [\partial_W(\mathcal{M}v)]_t(x) = (0, (\mathcal{M}v)_t(x))$ for $t \in [0,T)$.

Second step. Now we construct a solution on $[0,T]$ by a contraction argument when $T \leqslant 1$ (the same argument applies when $T \geqslant 1$ thanks to the remark made at the beginning of the first step). We choose λ large enough such that $C\kappa\exp(CT\vartheta^2)\lambda^{-\epsilon} \leqslant 1/4$ (with the same

constant C as in the statement of Theorem 12) and we remark that $(\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W),\Theta_T^{\vartheta,\lambda})$ is a Banach space. We then notice that, for any $u,v\in\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$, $\widehat{\mathcal{M}}u-\widehat{\mathcal{M}}v=\mathcal{M}(u-v)$ (the equality holding true for the lifted versions), so that, by Theorem 12 and Picard's theorem, the mapping $\widehat{\mathcal{M}}$ admits a unique fixed point $\bar v$ in $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$. Letting

$$\begin{aligned}\bar u_t(x)&=\int_{\mathbb{R}}p_{T-t}(x-y)u_T(y)\,\mathrm{d}y-\int_t^T\int_{\mathbb{R}}p_{s-t}(x-y)f_s(y)\,\mathrm{d}y\,\mathrm{d}s\\&\quad+\int_t^T\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y\bar v_s(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s,\end{aligned}\tag{36}$$

we obtain a mild solution, as defined in (12). It must be unique as the x-derivative of any other mild solution (when lifted up) is a fixed point of $\widehat{\mathcal{M}}$. Differentiation under the integral symbol in the mild formulation (12) can be justified by Lemma 11, making use of a standard uniform integrability argument.

Third step. We finally prove (13) and (14). We start with estimating $\bar v$. With our choice of λ and by Theorem 12, we have

$$\Theta_T^{\vartheta,\lambda}(\bar v)\leqslant\Theta_T^{\vartheta,\lambda}(\widehat{\mathcal{M}}0)+\frac{3}{4}\Theta_T^{\vartheta,\lambda}(\bar v),$$

where 0 stands for the null function, so that

$$\Theta_T^{\vartheta,\lambda}(\bar v)\leqslant 4\Theta_T^{\vartheta,\lambda}(\widehat{\mathcal{M}}0).\tag{37}$$

As $\widehat{\mathcal{M}}0$ has a zero derivative with respect to W, it is well checked that $\Theta_T^{\vartheta,\lambda}(\widehat{\mathcal{M}}0)\leqslant C(\sup_{r\geqslant1}e^{-\vartheta r}[\![(u^T)']\!]_\beta^{[-r,r]}+\|f\|_\infty)$ for a universal constant C (which would depend on T_0 if T was assumed to be less than T_0 for some $T_0\geqslant1$). This gives the exponential bound for $\bar v$ and for the $(\beta/2,\beta)$-Hölder constant of $\bar v$ in time and space.

In order to get the same estimate for $\bar u$, we go back to the original formulation (12):

$$\begin{aligned}u_t(x)&=\int_{\mathbb{R}}p_{T-t}(x-y)u_T(y)\,\mathrm{d}y-\int_t^T\int_{\mathbb{R}}p_{s-t}(x-y)f_s(y)\,\mathrm{d}y\,\mathrm{d}s\\&\quad+\int_t^T\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y\partial_x u_s(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s.\end{aligned}\tag{38}$$

Again, the two first terms can be estimated by standard properties of the heat kernel: the first term is at most of exponential growth and it is differentiable in time, the time derivative being also at most of exponential growth; the second term is bounded and it is γ-Hölder continuous on the whole space for any $\gamma\in(0,1)$. Finally the third term can be handled by repeating the analysis of $\mathcal{M}v$ in the proof of Theorem 12: Following (27) and (28), it is at most of exponential growth and it is locally $(1+\beta)/2$-Hölder continuous in time, the Hölder constant growing at most exponentially fast in the space variable (in comparison with (28), the additional $1/2$ comes from the fact there is one derivative less in the heat kernel).

3.3. **Proof of Proposition 6.** As above, we omit the superscript T in $Z^{n,T}$, $W^{n,T}$ and $\mathscr{W}^{n,T}$. Stability of solutions under mollification of the input follows from a classical compactness argument. Given a sequence $(W^n,\mathscr{W}^n)_{n\geqslant1}$ as in the statement, we can solve (12) for any

$n \geqslant 1$: The solution is denoted by u^n and its gradient by $v^n = \partial_x u^n$. By (2) in Proposition 6 and by the previous subsection, it is well-checked that

$$\sup_{n\geqslant 1} \Theta_T^{\vartheta,\lambda}(v^n) < \infty. \tag{39}$$

Therefore, by Theorem 12,

$$\sup_{n\geqslant 1} \Theta_T^{\vartheta,\lambda}(\mathcal{M}^n v^n) < \infty, \tag{40}$$

where $\mathcal{M}^n$ is obtained by replacing Y by Y^n in the definition of $\mathcal{M}$. It is worth mentioning that, contrary to the convention we have used so far, we must use $[\partial_{W^n}(\mathcal{M}^n v^n)]_t = (0, (\mathcal{M}^n v^n)_t)$ as choice of the derivative and not $[\partial_{W^n}(\mathcal{M}^n v^n)]_t = 0$ (which was the convention for smooth functions).

As a consequence of (39) and (40), we deduce that the sequences $(v^n)_{n\geqslant 1}$ and $(\mathcal{M}^n v^n)$ are uniformly continuous on compact subsets of $[0,T]\times\mathbb{R}$. In the same way, the sequence $(u^n)_{n\geqslant 1}$ is also uniformly continuous on compact subsets. Moreover, u^n, v^n and $\mathcal{M}^n v^n$ are at most of exponential growth, uniformly in $n \geqslant 1$. By Arzelà-Ascoli Theorem, we can extract subsequences (still indexed by n) that converge uniformly on compact subsets of $[0,T]\times\mathbb{R}$. Limits of $(u^n)_{n\geqslant 1}$, $(v^n)_{n\geqslant 1}$ and $(\mathcal{M}^n v^n)_{n\geqslant 1}$ are respectively denoted by $\hat{u}$, $\hat{v}$ and $\hat{m}$. In order to complete the proof, we must prove that $(\hat{u},\hat{v})$ is a mild solution of (12).

By (39), the sequence $(\mathscr{R}^{v_t^n})_{n\geqslant 1}$ is uniformly bounded on compact subsets of $[0,T)\times\mathbb{R}$. Writing (7) for each of the v^n and letting n tend to ∞, this says that the pair $(\hat{v},(0,\hat{m}))$ belongs to $\mathcal{B}^{\beta,\vartheta}([0,T),\mathbb{R})$, the remainder at any time $t\in[0,T)$ being denoted by $\hat{\mathscr{R}}^t$. At this stage of the proof, we know that, for any $t\in[0,T)$ and any $r\geqslant 1$, $[\![\hat{v}_t - v_t^n]\!]_\beta^{[-r,r]} + [\![\hat{m}_t - \partial_{W^n} v_t^n]\!]_\beta^{[-r,r]}$ tends to 0 as n tends to the infinity. We wish we also had $\lim_n [\![\hat{\mathscr{R}}^t - \mathscr{R}^{v_t^n}]\!]_{2\beta}^{[-r,r]} = 0$ in order to pass to the limit in the rough integrals involved in the mild formulation, as the convergence of the remainders is required to do so (see Subsection 2.1). Actually, we cannot prove it. Anyhow, by (7), the convergence holds in $L^\infty([-r,r])$ so that, by (39), it holds as well in Hölder norm, but with β replaced by any $\beta'<\beta$, that is $\lim_n [\![\hat{\mathscr{R}}^t - \mathscr{R}^{v_t^n}]\!]_{2\beta'}^{[-r,r]} = 0$.

Replacing β by β', we can pass to the limit in the rough integrals appearing in the mild formulation (12) of the PDE satisfied by each of the $(v^n)_{n\geqslant 1}$'s. To pass to the limit in the whole formulation, we can invoke some uniform integrability argument as we did to differentiate the mild formulation after Eq. (36). Thus the pair $(\hat{v},(0,\hat{m}))$ satisfies $\hat{v} = \widehat{\mathcal{M}}\hat{v}$ in $\mathcal{B}^{\beta',\vartheta}([0,T)\times\mathbb{R},W)$, which is enough to identify with the solution in $\mathcal{B}^{\beta,\vartheta}([0,T)\times\mathbb{R},W)$.

4. Stochastic Calculus for the Solution

In Theorem 8, we proved existence and uniqueness of a solution to the martingale problem associated with (1), but we said nothing about the dynamics of the solution. In this section, we answer to this question and give a sense to the formulation (4).

4.1. **Recovering the Brownian part.**

Equation (4) suggests that the dynamics of the solution to (1) indeed involves some Brownian part. The point we discuss here is thus twofold: (i) We recover in a quite *canonical* way the Brownian part in the dynamics of the solution; (ii) we discuss the structure of the remainder.

Theorem 13. *Under the assumption of Theorem 8, for any given initial condition x_0, we can find a probability measure (still denoted by $\mathbb{P}$) on the* enlarged *canonical space $\mathcal{C}([0,T_0],\mathbb{R}^2)$*

(endowed with the canonical filtration $(\mathcal{F}_t)_{0\leqslant t\leqslant T_0}$) such that, under $\mathbb{P}$, the canonical process, denoted by $(X_t, B_t)_{0\leqslant t\leqslant T_0}$, satisfies the followings:

(i) The law of $(X_t)_{0\leqslant t\leqslant T_0}$ under $\mathbb{P}$ is a solution to the martingale problem with x_0 as initial condition at time 0 and the law of $(B_t)_{0\leqslant t\leqslant T_0}$ under $\mathbb{P}$ is a Brownian motion.

(ii) For any $q \geqslant 1$ and any $\beta < \alpha$, there is a constant $C = C(\alpha, \beta, \chi, \kappa_{\alpha,\chi}(W, \mathscr{W}), q, T_0)$ such that, for any $0 \leqslant t \leqslant t + h \leqslant T_0$,

$$\mathbb{E}\big[\big|X_{t+h} - X_t - (B_{t+h} - B_t)\big|^q\big]^{\frac{1}{q}} \leqslant Ch^{(1+\beta)/2}. \tag{41}$$

(iii) For any $0 \leqslant t \leqslant t + h \leqslant T_0$,

$$\mathbb{E}\big[X_{t+h} - X_t|\mathcal{F}_t\big] = \mathfrak{b}(t, X_t, h) := u_t^{t+h}(X_t) - X_t, \tag{42}$$

where the mapping $u^{t+h} : [0, t+h]\times\mathbb{R} \ni (s, x) \mapsto u^{t+h}(s, x)$ is the mild solution of $\mathcal{P}(Y, 0, t+h)$ with $u_{t+h}^{t+h}(x) = x$ as terminal condition.

Proof. The point is to come back to the proof of the solvability of the martingale problem in Subsection 2.7. For free and with the same notations, we have the tightness of the family $(X_t^n, B_t)_{0\leqslant t\leqslant T_0}$, which is sufficient to extract a converging subsequence. The (weak) limit is the pair $(X_t, B_t)_{0\leqslant t\leqslant T_0}$ in (i). (Pay attention that we do not claim that the 'B' at the limit is the same as the 'B' in the regularized problems but, for convenience, we use the same letter.) We then repeat the proof of (17) which writes:

$$\begin{aligned} X_{t+h}^n - X_t^n &= \int_t^{t+h} \partial_x u_s^n(X_s^n)\,\mathrm{d}B_s + u_t^n(X_t^n) - u_{t+h}^n(X_t^n) \\ &= B_{t+h} - B_t + \int_t^{t+h} \big[\partial_x u_s^n(X_s^n) - 1\big]\,\mathrm{d}B_s + \big[u_t^n(X_t^n) - u_{t+h}^n(X_t^n)\big]. \end{aligned}$$

Repeating the analysis of the the third step in Subsection 2.7, we know that the third term in the right hand side satisfies the bound (41). The point is thus to prove that the second term also satisfies this bound. Recalling that $u_{t+h}^n(x) = x$, we notice that $\partial_x u_s^n(X_s^n) - 1 = \partial_x u_s^n(X_s^n) - \partial_x u_{t+h}^n(X_s^n)$. The bound then follows from the fact that $\partial_x u^n$ is locally $\beta/2$-Hölder continuous in time, the Hölder constant being at most of exponential growth, as ensured by Theorem 5. Letting n tend to ∞, we complete the proof of (ii).

The last assertion (iii) is easily checked for with X replaced by X^n and u^{t+h} replaced by u^n (and for sure with $\mathcal{F}_t$ replaced by the σ-field generated by $(X_s^n, B_s)_{0\leqslant s\leqslant t}$). It is quite standard to pass to the limit in n. □

4.2. **Expansion of the drift.** The next proposition gives a more explicit insight into the shape of the function $\mathfrak{b}$ in (42):

Proposition 14. *Given $T_0 > 0$, there exist a constant C and an exponent $\varepsilon > 0$ such that*

$$\begin{aligned} \mathfrak{b}(t, x, h) &= b(t, x, h) + O\big(h^{1+\varepsilon}\exp(2|x|)\big), \\ b(t, x, h) &= \int_t^{t+h}\int_{\mathbb{R}} \partial_x p_{s-t}(x - y)\big(Y_s(y) - Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s \\ &\quad + \int_t^{t+h}\int_{\mathbb{R}} \partial_x p_{s-t}(x - y)\int_x^y Z_s^{t+h}(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s, \end{aligned}$$

$O(\cdot)$ standing for the Landau notation (the underlying constant in the Landau notation being uniform in $0 \leqslant t \leqslant t + h \leqslant T_0$).

Remark 15. *The first term in the definition of $b(t,x,h)$ reads as a mollification (in x) of the gradient (in x) of $(Y_t(x))_{t\leqslant s\leqslant t+h,x\in\mathbb{R}}$ by means of the transition density of $(B_t)_{t\geqslant 0}$ (which is the martingale process driving the dynamics of X). It is (locally in x) of order $h^{1/2+\alpha/2}$. The second term reads as a correction in the mollification of $(Y_s(x))_{t\leqslant s\leqslant t+h,x\in\mathbb{R}}$. It keeps track of the rough path structure of $(Y_s(x))_{t\leqslant s\leqslant t+h,x\in\mathbb{R}}$. The proof right below shows that it is of order $h^{1/2+\alpha}$, thus proving that it can be 'hidden' in the remainder $O(h^{1+\epsilon})$ when $\alpha>1/2$. This requirement $\alpha>1/2$ fits the standard threshold in the rough paths theory above which Young's theory applies.*

Proof. From (12), we know that $u_t^{t+h}(x)$ expands as

$$u_t^{t+h}(x)=x+\int_t^{t+h}\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y v_s^{t+h}(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s,$$

where $v_s^{t+h}(y)=\partial_x u_s^{t+h}(y)$. Taking into account the terminal condition $v_{t+h}^{t+h}\equiv 1$ and following (35), the equation for v can be reformulated into $(\widehat{\mathcal{M}}v)_t(x)=1+(\mathcal{M}v)_t(x)$, with the same notations as in Theorem 12 (with $T=t+h$). By Theorem 12, $(v,\partial_{W^{t+h}}v)$ belongs to $\mathcal{B}^{\beta,\vartheta}([0,t+h)\times\mathbb{R},W^{t+h})$, with

$$\partial_Y\mathcal{M}(v,\partial_{W^{t+h}}v)_t(x)=0,\ \partial_{Z^{t+h}}\mathcal{M}(v,\partial_{W^{t+h}}v)_t(x)=v_t(x).$$

Therefore, we can write

$$v_s^{t+h}(z)=v_s^{t+h}(x)+v_s^{t+h}(x)\big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)+\mathscr{R}^{v_s}(x,z),$$

which we can plug into the expression for $u_t^{t+h}(x)$ by means of Theorem 1:

$$\begin{aligned}u_t^{t+h}(x)-x&=\int_t^{t+h}v_s^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\big(Y_s(y)-Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s\\&\quad+\int_t^{t+h}v_s^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y\big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s\\&\quad+\int_t^{t+h}\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\mathscr{U}_s^{t+h}(x,y)\,\mathrm{d}y\,\mathrm{d}s,\end{aligned}\tag{43}$$

where $\mathscr{U}_s^{t+h}(x,y)$ in a remainder term that derives from the approximation of the rough integral of v_s^{t+h} with respect to Y_s. By Theorem 1, there exist a constant C and an exponent $\varepsilon>0$ such that

$$\begin{aligned}&\left|\int_t^{t+h}\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\mathscr{U}_s^{t+h}(x,y)\,\mathrm{d}y\,\mathrm{d}s\right|\\&\leqslant C\exp(2|x|)\int_t^{t+h}(s-t)^{-\frac{1}{2}}\int_{\mathbb{R}}p_{s-t}(x-y)\exp(|x-y|)|x-y|^{1+\varepsilon}\,\mathrm{d}y\,\mathrm{d}s\\&\leqslant C\exp(2|x|)h^{1+\varepsilon}.\end{aligned}\tag{44}$$

Above, the exponential factor permits to handle the polynomial growth of $W^{t+h}=(Y,Z^{t+h})$ and the exponential growth of v^{t+h} (see the definition of $\Theta_T^{\vartheta,\lambda}(v)$ in the statement of Theorem 12), the exponent in the exponential factor being arbitrarily chosen as 1 (which leaves 'some space' to handle additional polynomial growth and which is possible since the terminal condition u_{t+h}^{t+h} is of polynomial growth).

We now investigate the second term in the right hand side of (43). We recall that, by assumption, there exists a constant C, independent of h, such that

$$\left|\int_x^y \big(Z_s^{t+h}(z) - Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\right| \leqslant \big|\mathscr{W}_s^{t+h}(x,y)\big| \leqslant C(1+|x|\vee|y|)^{2\chi}|x-y|^{2\alpha}. \tag{45}$$

We also recall from Theorem 5 that v is $(\alpha-\epsilon)/2$-Hölder continuous in time, locally in space (the rate of growth of the Hölder constant being at most exponential and Theorem 12 allowing to choose 1 as exponent in the exponential), so that $|v_s^{t+h}(y)-1| \leqslant Ch^{(\alpha-\epsilon)/2}\exp(|y|)$, for $s\in[t,t+h]$ and for a possibly new value of the constant C. Therefore,

$$\begin{aligned}
&\int_t^{t+h} v_s^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y \big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s\\
&=\int_t^{t+h}\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y \big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s\\
&\quad+\int_t^{t+h}\big(v_s^{t+h}(x)-1\big)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y \big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s,
\end{aligned}$$

the last term being less than

$$C\exp(2|x|)h^{(\alpha-\epsilon)/2}\int_t^{t+h}(s-t)^{-1/2+\alpha}\,\mathrm{d}r \leqslant C\exp(2|x|)h^{1/2+3\alpha/2-\epsilon} \leqslant C\exp(2|x|)h^{1+\epsilon}, \tag{46}$$

the last inequality holding true since α is strictly larger than $1/3$ and ϵ can be chosen arbitrarily small. Therefore, from (43), (44) and (45), we deduce that

$$\begin{aligned}
u_t^{t+h}(x)-x &= \int_t^{t+h} v_s^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\big(Y_s(y)-Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s\\
&\quad+\int_t^{t+h}\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\int_x^y \big(Z_s^{t+h}(z)-Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s+O\big(\exp(2|x|)h^{1+\epsilon}\big).
\end{aligned} \tag{47}$$

Using (45) once more and following the proof of (46), we also have

$$u_t^{t+h}(x)-x=\int_t^{t+h} v_s^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{s-t}(x-y)\big(Y_s(y)-Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s+O\big(\exp(2|x|)h^{1/2+\alpha}\big).$$

It then remains to look at the first term in the right-hand side of (43). The point is to expand $v_t^{t+h}(x)$ on the same model as $u_t^{t+h}(x)$ right above. Basically, the same expansion holds but, because of the derivative in the definition of $v_t^{t+h}(x)=\partial_x u_t^{t+h}(x)$, we loose $1/2$ in the power of h in the Landau notation. Therefore, for $t\leqslant s\leqslant t+h$, the above expansion turns into

$$v_s^{t+h}(x)-1=\int_s^{t+h} v_\rho^{t+h}(x)\int_{\mathbb{R}}\partial_x p_{\rho-t}(x-y)\big(Y_s(y)-Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}\rho+O\big(\exp(2|x|)h^{\alpha}\big).$$

Using once again the fact that v^{t+h} is $(\alpha-\epsilon)/2$-Hölder continuous in time (locally in space, the Hölder constant being at most of exponential growth), we obtain

$$\begin{aligned} v_s^{t+h}(x) - 1 &= \int_s^{t+h} \int_{\mathbb{R}} \partial_x^2 p_{\rho-t}(x-y)\big(Y_s(y) - Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}\rho \\ &\quad + \int_s^{t+h} \big(v_\rho^{t+h}(x) - 1\big) \int_{\mathbb{R}} \partial_x^2 p_{\rho-t}(x-y)\big(Y_s(y) - Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}\rho + O\big(\exp(2|x|)h^\alpha\big) \\ &= Z_s^{t+h}(x) + O\bigg(\exp(2|x|)\bigg[h^\alpha + h^{(\alpha-\epsilon)/2}\int_s^{t+h}(\rho-t)^{-1+\alpha/2}\,\mathrm{d}\rho\bigg]\bigg). \end{aligned}$$

The last term can be bounded by $O(\exp(2|x|)h^{\alpha-\epsilon/2})$. Now, by (47),

$$\begin{aligned} u_t^{t+h}(x) - x &= \int_t^{t+h} \big(1 + Z_s^{t+h}(x)\big) \int_{\mathbb{R}} \partial_x p_{s-t}(x-y)\big(Y_s(y) - Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s \\ &\quad + \int_t^{t+h} \int_{\mathbb{R}} \partial_x p_{s-t}(x-y) \int_x^y \big(Z_s^{t+h}(z) - Z_s^{t+h}(x)\big)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}r \\ &\quad + O\bigg(\exp(2|x|)\bigg[h^{\alpha-\epsilon/2}\int_t^{t+h}(s-t)^{-1/2+\alpha/2}ds + h^{1+\epsilon}\bigg]\bigg). \end{aligned} \tag{48}$$

It thus remains to bound

$$\int_t^{t+h} Z_s^{t+h}(x) \int_{\mathbb{R}} \partial_x p_{s-t}(x-y)\big(Y_s(y) - Y_s(x)\big)\,\mathrm{d}y\,\mathrm{d}s.$$

By (11), it is plain to see that $Z_r^{t+h}(x) = O(\exp(2|x|)h^{\alpha/2})$. Then, the above term must at most of order $O(\exp(2|x|)h^{1/2+\alpha})$, from which the proof of the proposition is easily completed.

In order to complete the proof of Remark 15, it remains to show the announced bound for

$$\int_t^{t+h} \int_{\mathbb{R}} \partial_x p_{s-t}(x-y) \int_x^y Z_s^{t+h}(z)\,\mathrm{d}Y_s(z)\,\mathrm{d}y\,\mathrm{d}s.$$

We already have a bound when $Z_s^{t+h}(z)$ is replaced by $Z_s^{t+h}(x)$. By (45), we also have a bound when $Z_r^{t+h}(z)$ is replaced by $Z_r^{t+h}(z) - Z_r^{t+h}(x)$. □

4.3. **Purpose.** The goal is now to prove that Theorem 13 and Proposition 14 are sufficient to define a differential calculus for which the infinitesimal variation dX_t reads

$$\mathrm{d}X_t = \mathrm{d}B_t + b(t, X_t, \mathrm{d}t), \quad t \in [0,T), \tag{49}$$

or, in a macroscopic way, $X_t = X_0 + B_t + \int_0^t b(s, X_s, \mathrm{d}s)$, which gives a sense to (1). In that framework, Proposition 14 and Remark 15 give some insight into the shape of the drift.

As explained below, we are able to define a stochastic calculus in such a way that the process $(\int_0^t b(s, X_s, \mathrm{d}s))_{0\leqslant t\leqslant T}$ has a Hölder continuous version, with $(1+\alpha)/2 - \epsilon$ as Hölder exponent, for $\epsilon > 0$ as small as desired, thus making $(X_t)_{0\leqslant t\leqslant T}$ a Dirichlet process.

In order to give a meaning to (49), the point is to give a sense to $\int_0^T \psi_t\,\mathrm{d}X_t$ and possibly to $\int_0^T \psi_t b(t, X_t, \mathrm{d}t)$ for a sufficiently large class of integrands: We construct the integral with respect to processes $(\psi_t)_{0\leqslant t\leqslant T}$ that are progressively-measurable and $(1-\alpha)/2 + \epsilon$ Hölder continuous in L^p for some $p > 2$ and some $\epsilon > 0$. The construction of the integral consists of a mixture of Young's and Itô's integrals. Precisely, the progressive-measurability of $(\psi_t)_{0\leqslant t\leqslant T}$ permits to 'get rid of' the martingale increments in X that are different from the Brownian

ones and thus to focus on the function b only in order to define the non-Brownian part of the dynamics. Then, the Hölder property of $(\psi_t)_{0\leqslant t\leqslant T}$ permits to integrate with respect to $(b(t,X_t,\mathrm{d}t))_{0\leqslant t\leqslant T}$ in a Young sense. For that reason, the resulting integral is called a stochastic Young integral. It is worth mentioning that it permits to consider within the same framework integrals defined with respect to the martingale part of X and integrals defined with respect to the zero quadratic variation part of X.

The construction we provide below is given in a larger set-up. In the whole section, we thus use the following notation: $(\Omega,(\mathcal{F}_t)_{t\geqslant 0},\mathbb{P})$ denotes a filtered probability space satisfying the usual conditions; moreover, for any $0\leqslant s\leqslant t$, $\mathcal{S}(s,t)$ denotes the set $\{s'\in[0,s],t'\in[0,t],s'\leqslant t'\}$. The application to (41) is discussed in Subsection 4.6.

4.4. L^p Construction of the Integral.

4.4.1. *Materials.* We are given a real $T>0$ and a continuous progressively-measurable process $(A(s,t))_{0\leqslant s\leqslant t\leqslant T}$ in the sense that, for any $0\leqslant s\leqslant t$, the mapping $\Omega\times\mathcal{S}(s,t)\ni(\omega,s',t')\mapsto A(s',t')$ is measurable for the product σ-field $\mathcal{F}_t\otimes\mathcal{B}(\mathcal{S}(s,t))$ and the mapping $\mathcal{S}(T,T)\ni(s,t)\mapsto A(s,t)$ is continuous. We assume that there exist a constant $\Gamma\geqslant 0$, three exponents $\varepsilon_0\in(0,1/2]$, $\varepsilon_1,\varepsilon_1'>0$ and a real $q\geqslant 1$ such that, for any $0\leqslant t\leqslant t+h\leqslant t+h'\leqslant T$,

$$
\begin{aligned}
&\mathbb{E}\big[\big|\mathbb{E}\big[A(t,t+h)|\mathcal{F}_t\big]\big|^q\big]^{\frac{1}{q}}\leqslant\Gamma h^{\frac{1}{2}+\varepsilon_0},\\
&\mathbb{E}\big[|A(t,t+h)|^q\big]^{\frac{1}{q}}\leqslant\Gamma h^{\frac{1}{2}},\\
&\mathbb{E}\big[\big|\mathbb{E}\big[A(t,t+h)+A(t+h,t+h')-A(t,t+h')|\mathcal{F}_t\big]\big|^q\big]^{\frac{1}{q}}\leqslant\Gamma(h')^{1+\varepsilon_1},\\
&\mathbb{E}\big[|A(t,t+h)+A(t+h,t+h')-A(t,t+h')|^q\big]^{\frac{1}{q}}\leqslant\Gamma(h')^{\frac{1}{2}(1+\varepsilon_1')}.
\end{aligned}
\tag{50}
$$

In the framework of (49), we have in mind to choose $A(t,t+h)=X_{t+h}-X_t$ or $A(t,t+h)=B_{t+h}-B_t$, in which cases A has an additive structure and ε_1 and ε_1' can be chosen as large as desired, or $A(t,t+h)=b(t,X_t,h)$, in which case A is not additive. The precise application to (49) is detailed in Subsection 4.6. Generally speaking, we call $A(t,t+h)$ a *pseudo-increment*. Considering pseudo-increments instead of increments (that enjoy, in comparison with, an additive property) allows more flexibility and permits, as just said, to give a precise meaning to $b(t,X_t,\mathrm{d}t)$ in (49). The strategy is then to split $A(t,t+h)$ into two pieces:

$$
R(t,t+h):=\mathbb{E}\big[A(t,t+h)|\mathcal{F}_t\big],\quad M(t,t+h):=A(t,t+h)-\mathbb{E}\big[A(t,t+h)|\mathcal{F}_t\big],
\tag{51}
$$

$M(t,t+h)$ being understood as a sort of martingale increment and $R(t,t+h)$ as a sort of drift.

We are also given a continuous progressively-measurable process $(\psi_t)_{0\leqslant t\leqslant T}$ and we assume that, for an exponent $\varepsilon_2<\varepsilon_0$ and for any $0\leqslant t\leqslant t+h\leqslant T$,

$$
\mathbb{E}\big[|\psi_t|^{q'}\big]^{\frac{1}{q'}}\leqslant\Gamma,\quad\mathbb{E}\big[\psi_{t+h}-\psi_t|^{q'}\big]^{\frac{1}{q'}}\leqslant\Gamma h^{\frac{1}{2}-\varepsilon_2},
\tag{52}
$$

for some $q'\geqslant 1$. We then let $p=qq'/(q+q')$ so that $1/p=1/q+1/q'$.

4.4.2. *Objective.* The aim of the subsection is to define the stochastic integral $\int_0^T \psi_t A(t,t+\mathrm{d}t)$ as an $L^p(\Omega,\mathbb{P})$ version of the Young integral. In comparison with the standard version of the Young integral, the $L^p(\Omega,\mathbb{P})$ construction will benefit from the martingale structure of the pseudo-increments $(M(t,t+h))_{0\leqslant t\leqslant t+h\leqslant T}$, the integral being defined as the $L^p(\Omega,\mathbb{P})$ limit of Riemann sums as the step size of the underlying subdivision tends to 0. Given a subdivision $\Delta=\{0=t_0<t_1<\cdots<t_N=T\}$, we thus define the Δ-Riemann sum

$$S(\Delta):=\sum_{i=0}^{N-1}\psi_{t_i}A(t_i,t_{i+1}). \tag{53}$$

We emphasize that this definition is exactly the same as the one used to define Itô's integral: on the step $[t_i,t_{i+1}]$, the process ψ is approximated by the value at the initial point t_i. For that reason, we will say that the Riemann sum is *adapted*. In that framework, we claim:

Theorem 16. *There exists a constant $C=C(q,q',\Gamma,\varepsilon_0,\varepsilon_1,\varepsilon_2)$, such that, given two subdivisions $\Delta\subset\Delta'$, with $\pi(\Delta)\leqslant 1$,*

$$\mathbb{E}\big[|S(\Delta)-S(\Delta')|^p\big]^{1/p}\leqslant C'\max(T^{1/2},T)\big(\pi(\Delta)\big)^{\eta}, \tag{54}$$

where $\pi(\Delta)$ denotes the step size of the subdivision Δ, that is $\pi(\Delta):=\max_{1\leqslant i\leqslant N}[t_i-t_{i-1}]$, and with $\eta:=\min(\varepsilon_0-\varepsilon_2,\varepsilon_1,\varepsilon_1'/2)$.

For general partitions Δ and Δ' (without any inclusion requirement), Theorem 16 applies to the pairs $(\Delta,\Delta\cup\Delta')$ and $(\Delta',\Delta\cup\Delta')$, so that (54) holds in that case as well provided $\pi(\Delta)$ in the right-hand side is replaced by $\max(\pi(\Delta),\pi(\Delta'))$. We deduce that $S(\Delta)$ has a limit in $L^p(\Omega,\mathbb{P})$ as $\pi(\Delta)$ tends to 0. We call it the stochastic Young integral of ψ with respect to the pseudo-increments of A.

4.4.3. *Proof of Theorem 16. First Step.* First, we consider the case where the two subdivisions Δ and Δ', Δ being included in Δ', are not so different one from each other. Precisely, given $\Delta=\{0=t_0<t_1<\cdots<t_N=T\}$ and $\Delta'=\Delta\cup\{t_1'<\cdots<t_L'\}$ ($L\geqslant 1$), the $(t_i)_{1\leqslant i\leqslant N}$'s and the $(t_j')_{1\leqslant j\leqslant L}$'s being pairwise distinct, we assume that, between two consecutive points in Δ, there is at most one point in Δ'. For any $j\in\{1,\dots,L\}$, we then denote by s_j^- and s_j^+ the largest and smallest points in Δ such that $s_j^-<t_j'<s_j^+$. We have $t_j'<s_j^+\leqslant s_{j+1}^-<t_{j+1}'$ for $1\leqslant j\leqslant L-1$. We then claim:

Lemma 17. *Under the above assumption, the estimate* (54) *holds with $\pi(\Delta)$ replaced by $\rho(\Delta'\backslash\Delta)$, where $\rho(\Delta'\backslash\Delta):=\sup_{1\leqslant j\leqslant L}[s_j^+-s_j^-]$.*

Proof of Lemma 17. (i) As a first step, we compute the difference $S(\Delta')-S(\Delta)$. We write

$$S(\Delta')-S(\Delta)=\sum_{j=1}^{L}\big[S(\Delta^j)-S(\Delta^{j-1})\big],$$

with $\Delta^j=\Delta\cup\{t_1',\dots,t_j'\}$, for $1\leqslant j\leqslant L$, and $\Delta^0=\Delta$. Then,

$$\begin{aligned}S(\Delta^j)&=S(\Delta^{j-1})+\psi_{s_j^-}A(s_j^-,t_j')+\psi_{t_j'}A(t_j',s_j^+)-\psi_{s_j^-}A(s_j^-,s_j^+)\\&=S(\Delta^{j-1})+\big(\psi_{t_j'}-\psi_{s_j^-}\big)A(t_j',s_j^+)+\psi_{s_j^-}\big(A(s_j^-,t_j')+A(t_j',s_j^+)-A(s_j^-,s_j^+)\big).\end{aligned}$$

Therefore,

$$
\begin{aligned}
S(\Delta') - S(\Delta) &= \sum_{j=1}^{L} \big(\psi_{t'_j} - \psi_{s_j^-}\big) M(t'_j, s_j^+) + \sum_{j=1}^{L} \big(\psi_{t'_j} - \psi_{s_j^-}\big) R(t'_j, s_j^+) \\
&\quad + \sum_{j=1}^{L} \psi_{s_j^-} \big(A(s_j^-, t'_j) + A(t'_j, s_j^+) - A(s_j^-, s_j^+)\big) \\
&:= \delta_1 S(\Delta, \Delta', M) + \delta_1 S(\Delta, \Delta', R) + \delta_2 S(\Delta, \Delta').
\end{aligned} \tag{55}
$$

(ii) We first investigate $\delta_1 S(\Delta, \Delta', M)$. The process $(\sum_{j=1}^{\ell} (\psi_{t'_j} - \psi_{s_j^-}) M(t'_j, s_j^+))_{0 \leqslant \ell \leqslant L}$ is a discrete stochastic integral and thus a martingale with respect to the filtration $(\mathcal{F}_{s_\ell^+})_{0 \leqslant \ell \leqslant L}$, with the convention that $s_0^- = s_0^+ = 0$. The sum of the squares of the increments is given by $\sum_{j=1}^{L} (\psi_{t'_j} - \psi_{s_j^-})^2 (M(t'_j, s_j^+))^2$. By the second line in (50) and by (52), we observe from Minkowski's inequality first and then from Hölder's inequality (recalling $1/p = 1/q + 1/q'$) that there exists a constant C such that

$$
\begin{aligned}
\mathbb{E}\bigg[\bigg|\sum_{j=1}^{L} \big(\psi_{t'_j} - \psi_{s_j^-}\big)^2 \big(M(t'_j, s_j^+)\big)^2\bigg|^{\frac{p}{2}}\bigg]^{\frac{2}{p}} &\leqslant \sum_{j=1}^{L} \mathbb{E}\Big[\big(\psi_{t'_j} - \psi_{s_j^-}\big)^p \mathbb{E}\Big[\big(M(t'_j, s_j^+)\big)^p | \mathcal{F}_{t'_j}\Big]\Big]^{\frac{2}{p}} \\
&\leqslant C \sum_{j=1}^{L} \big(t'_j - s_j^-\big)^{(1-2\varepsilon_2)} \big(s_j^+ - t'_j\big) \leqslant CT\big(\rho(\Delta' \backslash \Delta)\big)^{\eta_1},
\end{aligned}
$$

with $\eta_1 := 1 - 2\varepsilon_2 \geqslant 2(\varepsilon_0 - \varepsilon_2)$, where we have used $s_j^- < t'_j < s_j^+$. By discrete Burkholder-Davis-Gundy inequalities, we deduce that $\mathbb{E}[|\delta_1 S(\Delta, \Delta', M)|^p]^{1/p} \leqslant CT^{1/2} (\rho(\Delta' \backslash \Delta))^{\eta_1/2}$.

(iii) We now turn to $\delta_1 S(\Delta, \Delta', R)$. In the same way, by the first line in (50) and by (52),

$$
\begin{aligned}
\mathbb{E}\big[\big|\delta_1 S(\Delta, \Delta', R)\big|^p\big]^{\frac{1}{p}} &\leqslant \sum_{j=1}^{L} \mathbb{E}\big[\big||\psi_{t'_j} - \psi_{s_j^-}|^p |R(t'_j, s_j^+)|^p\big]^{\frac{1}{p}} \\
&\leqslant C \sum_{j=1}^{L} \big(t'_j - s_j^-\big)^{1/2-\varepsilon_2} \big(s_j^+ - t'_j\big)^{1/2+\varepsilon_0} \leqslant CT\big(\rho(\Delta' \backslash \Delta)\big)^{\eta_2},
\end{aligned}
$$

with $\eta_2 := \varepsilon_0 - \varepsilon_2$. Therefore, $\mathbb{E}[|\delta_1 S(\Delta, \Delta', R)|^p]^{1/p} \leqslant CT\big(\rho(\Delta' \backslash \Delta)\big)^{\eta_2}$.

(iv) We finally investigate $\delta_2 S(\Delta, \Delta')$. We split it into two pieces:

$$
\begin{aligned}
\delta_2 S(\Delta, \Delta') &= \sum_{j=1}^{L} \psi_{s_j^-} R'(s_j^-, t'_j, s_j^+) + \sum_{j=1}^{L} \psi_{s_j^-} M'(s_j^-, t'_j, s_j^+), \\
&:= \delta_2 S(\Delta, \Delta', R') + \delta_2 S(\Delta, \Delta', M'),
\end{aligned} \tag{56}
$$

with

$$
\begin{aligned}
R'(s_j^-, t'_j, s_j^+) &:= \mathbb{E}\big[A(s_j^-, t'_j) + A(t'_j, s_j^+) - A(s_j^-, s_j^+) \big| \mathcal{F}_{s_j^-}\big], \\
M'(s_j^-, t'_j, s_j^+) &:= A(s_j^-, t'_j) + A(t'_j, s_j^+) - A(s_j^-, s_j^+) - R'(s_j^-, t'_j, s_j^+).
\end{aligned}
$$

By the third line in (50) and by (52), we have, with $\eta_3 := \varepsilon_1$, $\mathbb{E}[|\delta_2 S(\Delta, \Delta', R')|^p]^{1/p} \leqslant CT(\rho(\Delta' \backslash \Delta))^{\eta_3}$.

We finally tackle $\delta_2 S(\Delta,\Delta',M')$. We notice that it generates a discrete time martingale with respect to the filtration $(\mathcal{F}_{s_\ell^+})_{0\leqslant\ell\leqslant L}$. As in the second step, we compute the $L^{p/2}(\Omega,\mathbb{P})$ norm of the sum of the squares of the increments. By the last line in (50), it is given by

$$\mathbb{E}\Big[\Big|\sum_{j=1}^{L}\psi^2_{s_j^-}\big(M'(s_j^-,t'_j,s_j^+)\big)^2\Big|^{\frac{p}{2}}\Big]^{\frac{2}{p}}\leqslant\sum_{j=1}^{L}\mathbb{E}\Big[\psi^p_{s_j^-}\mathbb{E}\Big[\big(M'(s_j^-,t'_j,s_j^+)\big)^p|\mathcal{F}_{s_{j-1}^+}\Big]\Big]^{\frac{2}{p}}\leqslant CT\big(\rho(\Delta'\backslash\Delta)\big)^{\eta_4},$$

with $\eta_4:=\varepsilon'_1$. By discrete Burkholder-Davis-Gundy inequality, $\mathbb{E}[|\delta_2 S(\Delta,\Delta',M')|^p]^{1/p}\leqslant CT^{1/2}(\rho(\Delta'\backslash\Delta))^{\eta_4/2}$.

Putting (i), (ii), (iii) and (iv) together, this completes the proof. □

4.4.4. *Proof of Theorem 16. Second Step.* We now consider the general case when $\Delta\subset\Delta'$ ($\Delta'\neq\Delta$) without any further assumption on the difference $\Delta'\backslash\Delta$.

As above, we denote the points in Δ by $t_1,\dots,t_N$. The points in the difference $\Delta'\backslash\Delta$ are denoted in the following way. For $i=1,\dots,N$, we denote by $t'_{1,i},\dots,t'_{L_i,i}$ the points in the intersection $(\Delta'\backslash\Delta)\cap(t_{i-1},t_i)$, where L_i denotes the number of points in $(\Delta'\backslash\Delta)\cap(t_{i-1},t_i)$. Each L_i may be written as $L_i=2\ell_i+\varepsilon_i$ where $\ell_i\in\mathbb{N}$ and $\varepsilon_i\in\{0,1\}$. We then define Δ'_1 as the subdivision made of the points that are in Δ together with the points

$$\big\{\{t'_{2\ell,i},\ \ell=1,\dots,\ell_i\}\cup\{t_{2\ell_i+1}\text{ if }\varepsilon_i=1\}\big\}\quad\text{whenever }\ell_i\geqslant1,\quad\text{for } i=1,\dots,N.$$

This says that, to construct Δ'_1, we delete, for any $i=1,\dots,N$, the point $t'_{1,i}$ if $L_i=1$ and the points that are in $(\Delta'\backslash\Delta)\cap(t_{i-1},t_i)$ and that have an odd index $2\ell-1$ with $1\leqslant\ell\leqslant\ell_i$ if $L_i>1$ (so that the last point is kept even if labelled by an odd integer when $\ell_i\geqslant1$). By construction, Δ'_1 and Δ' satisfy the assumption of Subsection 4.4.3, so that

$$\big\|S(\Delta'_1)-S(\Delta')\big\|_{L^p(\Omega,P)}\leqslant C\max(T^{1/2},T)\big[\rho(\Delta'\backslash\Delta'_1)\big]^{\eta}.$$

It holds $\Delta'_1\supset\Delta$. If $\Delta'_1\neq\Delta$, we then build a new subdivision Δ'_2 as the subdivision associated with Δ'_1 in the same manner as Δ'_1 is associated with Δ'. We then obtain

$$\big\|S(\Delta'_2)-S(\Delta'_1)\big\|_{L^p(\Omega,\mathbb{P})}\leqslant C\max(T^{1/2},T)\big[\rho(\Delta'_1\backslash\Delta'_2)\big]^{\eta}.\tag{57}$$

We then carry on the construction up until we reach $\Delta'_M=\Delta$ for some integer $M\geqslant1$. We notice that such an M does exist: by construction each Δ'_j contains Δ and $\sharp[\Delta'_j]<\sharp[\Delta'_{j-1}]$ (with the convention $\Delta'_0=\Delta'$).

We now make an additional assumption: We assume that Δ' is a dyadic subdivision, that is $\Delta'=\{2^{-P}kT,0\leqslant k\leqslant2^P\}$ for some $P\geqslant1$. This says that Δ is also made of dyadic points of order P. We denote by Q the unique integer such that

$$\max(L_i,1\leqslant i\leqslant N)=2^Q+r\quad\text{with }0\leqslant r\leqslant2^Q-1,$$

and by i_Q some index such that $L_{i_Q}=2^Q+r$. At the first step, the 2^Q first points in $(\Delta'\backslash\Delta)\cap(t_{i_Q-1},t_{i_Q})$ are reduced into 2^{Q-1} points. At the second step, they are reduced into 2^{Q-2} points and so on... Therefore, it takes steps to reduce the 2^Q first points in $(\Delta'\backslash\Delta)\cap(t_{i_Q-1},t_{i_Q})$ into a single one. Meanwhile, it takes at most Q steps to reduce the r remaining points in $(\Delta'\backslash\Delta)\cap(t_{i_Q-1},t_{i_Q})$ into a single one (without any interferences between the two reductions). We deduce that, after the Qth step, there are at most two operations

to perform to reduce Δ'_Q into Δ. This says that M is either $Q+1$ or $Q+2$ and that, at each step $j \in \{1,\dots,Q\}$ of the induction, we are doubling the step size $\rho(\Delta'_{j-1}\backslash\Delta'_j)$, that is

$$\rho(\Delta'_{j-1}\backslash\Delta'_j) = 2^{j-1}\rho(\Delta'\backslash\Delta'_1), \quad j=1,\dots,Q,$$

so that

$$\rho(\Delta'\backslash\Delta'_1) \leqslant 2^{-(Q-1)}\pi(\Delta), \quad \text{and} \quad \rho(\Delta'_{j-1}\backslash\Delta'_j) \leqslant 2^{j-Q}\pi(\Delta), \quad j=1,\dots,Q.$$

Therefore, $\rho(\Delta'_{j-1}\backslash\Delta'_j) \leqslant 2^{j-M+2}\pi(\Delta)$, $j=1,\dots,M$. By extending (57) to each of the steps of the induction, we get (up to a new value of C)

$$\|S(\Delta')-S(\Delta)\|_{L^p(\Omega,\mathbb{P})} \leqslant C\max(T^{1/2},T)\big[\pi(\Delta)\big]^\eta \sum_{j=0}^{M} 2^{\eta(j-M)} \leqslant C\max(T^{1/2},T)\big[\pi(\Delta)\big]^\eta. \tag{58}$$

When Δ and Δ' contain non-dyadic points (so that they are different from $\{0,T\}$), we can argue as follows. We can find a dyadic subdivision, denoted by D_2, such that, in any open interval delimited by two consecutive points in D_2, there is at most one element of Δ. Then, we remove points from D_2 to obtain a minimal subdivision D_1, made of dyadic points, such that, in any open interval delimited by two consecutive points in D_1, there is exactly one element of Δ. In such way, in any open interval delimited by two consecutive points in Δ, there is at most one point in D_1. Therefore, we can apply Lemma 17 to $(D_1, D_1\cup\Delta)$ and $(\Delta, D_1\cup\Delta)$. We get

$$\|S(D_1)-S(\Delta)\|_{L^p(\Omega,\mathbb{P})} \leqslant C\max(T^{1/2},T)\big[\max\big(\pi(D_1),\pi(\Delta)\big)\big]^\eta \leqslant C'\max(T^{1/2},T)\big[\pi(\Delta)\big]^\eta,$$

since $\pi(D_1) \leqslant 2\pi(\Delta)$. By the same argument, we can find a dyadic subdivision D'_1 for which the above inequality applies with (D_1,Δ) replaced by (D'_1,Δ'). Then, we can find a dyadic subdivision D such that both $D_1\subset D$ and $D'_1\subset D$. Applying (58) to (D_1,D) and to (D'_1,D), we can bound the difference between $S(D'_1)$ and $S(D_1)$. The result follows.

4.5. **Further Properties of the Integral.**

4.5.1. *Extension of the Integral.* Given the decomposition (51), it is worth noting that both the integrals $\int_0^T \psi_t M(t,t+dt)$ and $\int_0^T \psi_t R(t,t+dt)$ are also defined as L^p limits of the associated adapted Riemann sums. The main point is to check that Lemma 17 applies to S_M and S_R, where, with the same notation as in (53), $S_M(\Delta)=\sum_{i=0}^{N-1}\psi_{t_i}M(t_i,t_{i+1})$ and $S_R(\Delta)=\sum_{i=0}^{N-1}\psi_{t_i}R(t_i,t_{i+1})$. A careful inspection of the proof of Lemma 17 shows that the non-trivial point is to control the quantities $\delta_2 S(\Delta,\Delta',M)$ and $\delta_2 S(\Delta,\Delta',R)$, obtained by replacing A by M and R respectively in the definition of $\delta_2 S(\Delta,\Delta')$ in (55). Actually, since we already have a control of the sum of the two terms (as it coincides with $\delta_2 S(\Delta,\Delta')$ in the proof of Lemma 17), it is sufficient to control one of them only. Clearly,

$$\begin{aligned}\|\delta_2 S(\Delta,\Delta',R)\|_{L^p(\Omega,\mathbb{P})} \leqslant{}& \Big\|\sum_{j=1}^{L}\psi_{s_j^-}\big(R(s_j^-,t'_j)+\mathbb{E}\big(R(t'_j,s_j^+)|\mathcal{F}_{s_j^-}\big)-R(s_j^-,s_j^+)\big)\Big\|_{L^p(\Omega,\mathbb{P})}\\ &+\Big\|\sum_{j=1}^{L}\psi_{s_j^-}\big(R(t'_j,s_j^+)-\mathbb{E}\big(R(t'_j,s_j^+)|\mathcal{F}_{s_j^-}\big)\big)\Big\|_{L^p(\Omega,\mathbb{P})}.\end{aligned}$$

We emphasize that the first term above is nothing but $\delta_2 S(\Delta,\Delta',R')$ in (56), for which we already have a bound. Therefore, the only remaining point is to control the second term

above. Again, we notice that it has a martingale structure, which can be estimated by Burkholder-Davis-Gundy inequality. By the first line in (50) and by (52),

$$\mathbb{E}\Big[\Big|\sum_{j=1}^{L}\psi^2_{s_j^-}\Big(R(t'_j,s_j^+)-\mathbb{E}\big(R(t'_j,s_j^+)|\mathcal{F}_{s_j^-}\big)\Big)^2\Big|^{\frac{p}{2}}\Big]^{\frac{2}{p}}$$
$$\leqslant C\sum_{j=1}^{L}\mathbb{E}\Big[\psi^p_{s_j^-}\big(R(t'_j,s_j^+)\big)^p\Big]^{\frac{2}{p}}\leqslant C'\sum_{j=1}^{L}\big(s_j^+-s_j^-\big)^{1+2\varepsilon_0}\leqslant C''T\big(\rho(\Delta'\backslash\Delta)\big)^{2\varepsilon_0},$$

which is enough to conclude that Theorem 16 is also valid when replacing A by R or M in §4.4.4. Therefore, we are allowed to split the integral of ψ as $\int_0^T\psi_tA(t,t+\mathrm{d}t)=\int_0^T\psi_tM(t,t+\mathrm{d}t)+\int_0^T\psi_tR(t,t+\mathrm{d}t)$. The reader must pay attention to the fact that neither M nor R must satisfy (50) even if A does. The extension of the integral to the case when they are driven by M or R is thus a consequence of the proof of Theorem 16 itself.

4.5.2. *Continuity in Time.* It is plain to see that the integral is additive in the sense that, for any $0\leqslant S\leqslant S+S'\leqslant T$,

$$\int_0^{S+S'}\psi_tA(t,t+\mathrm{d}t)=\int_0^{S}\psi_tA(t,t+\mathrm{d}t)+\int_S^{S+S'}\psi_tA(t,t+\mathrm{d}t).$$

An important question in practice is the regularity property of the process $[0,T)\ni t\mapsto\int_0^t\psi_sA(s,s+\mathrm{d}s)$, which is not well-defined for the moment. At this stage of the procedure, each of the integrals is uniquely defined up to an event of zero probability which depends on t. A continuity argument is thus needed in order to give a sense to all the integrals at the same time. By Theorem 16, we know that, for $h\in(0,1)$,

$$\Big\|\int_t^{t+h}\psi_sA(s,s+\mathrm{d}s)-\psi_tA(t,t+h)\Big\|_{L^p(\Omega,\mathbb{P})}\leqslant Ch^{\frac{1}{2}+\eta},\tag{59}$$

for $\eta>0$ as in the statement of Theorem 16, so that, by the two first lines in (50), $\|\int_t^{t+h}\psi_sA(s,s+\mathrm{d}s)\|_{L^p(\Omega,\mathbb{P})}\leqslant Ch^{1/2}$, for possibly new values of C. By Kolmogorov's continuity criterion, this says that there exists a Hölder continuous version of the process $(\int_0^t\psi_sA(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T}$, with $1/2-1/p-\epsilon$ as pathwise Hölder exponent, for any $\epsilon>0$.

By the same argument, we notice that there exist Hölder continuous versions of the processes $(\int_0^t\psi_sM(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T}$ and $(\int_0^t\psi_sR(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T}$. The Hölder exponent of the second one is actually better. Indeed, noticing that (59) also holds for R and taking advantage of the first line in (50), we deduce that $\|\int_t^{t+h}\psi_sR(s,s+\mathrm{d}s)\|_{L^p(\Omega,\mathbb{P})}\leqslant Ch^{(1+\eta)/2}$, so that the pathwise Hölder exponent can be chosen as $(1+\eta)/2-1/p-\epsilon$ for any $\epsilon>0$.

4.5.3. *Dirichlet decomposition.* It is well-checked that the process $(\int_0^t\psi_sM(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T}$ is a martingale, thus showing that the integral of ψ with respect to the pseudo-increments of A can be split into two terms: a martingale and a drift. We expect that, in practical cases, the exponent p can be choose as large as desired: In this setting, the martingale part has $(1/2-\epsilon)$-Hölder continuous paths, for $\epsilon>0$ as small as desired, and the drift part has $(1/2+\eta-\epsilon)$-Hölder continuous paths, also for $\epsilon>0$ as small as desired, thus proving that the integral is a Dirichlet process.

4.6. **Application to diffusion processes driven by a distributional drift.** We now explain how the stochastic Young integral applies to (1). First, we can choose $A(t,t+h) = X_{t+h} - X_t$, for $0 \leqslant t \leqslant t+h \leqslant T_0$. Then the process A is additive. In particular, the two last lines in (50) are automatically satisfied with ε_1 and ε_1' as large as needed. By (41), the second line in (50) is also satisfied. Finally, we notice that

$$\mathbb{E}\big[X_{t+h} - X_t|\mathcal{F}_t\big] = \mathbb{E}\big[X_{t+h} - X_t - \big(B_{t+h} - B_t\big)|\mathcal{F}_t\big],$$

so that, by (41) again, the first line in (50) is satisfied with $\varepsilon_0 = \beta/2$.

With our construction, this permits to define $(\int_0^t \psi_s \,\mathrm{d}X_s)_{0\leqslant t\leqslant T_0}$ for any progressively measurable process $(\psi_t)_{0\leqslant t\leqslant T_0}$ satisfying (52) with $\varepsilon_2 < \beta/2$. It also permits to define the integrals $(\int_0^t \psi_s M(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ and $(\int_0^t \psi_s R(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T_0}$, where

$$M(t,t+h) = X_{t+h} - X_t - \mathbb{E}\big[X_{t+h} - X_t|\mathcal{F}_t\big], \quad R(t,t+h) = \mathbb{E}\big[X_{t+h} - X_t|\mathcal{F}_t\big].$$

By (42), we have $R(t,t+h) = \mathfrak{b}(t,X_t,h)$, so that $(\int_0^t \psi_s \mathfrak{b}(s,X_s,\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ is well-defined.

Moreover, by Proposition 14 and by boundedness of the exponential moments of $(X_t)_{0\leqslant t\leqslant T_0}$ (see the proof of Theorem 8), we know that $\hat{R}(t,t+h) = (b-\mathfrak{b})(t,X_t,h)$ also satisfies (50), from which we deduce that $(\int_0^t \psi_s (b-\mathfrak{b})(s,X_s,\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ and so $(\int_0^t \psi_s b(s,X_s,\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ are well-defined. Actually the exponent in the power of h appearing in the difference $(b-\mathfrak{b})(t,X_t,h)$ being strictly greater than 1, the integral process $(\int_0^t \psi_s (b-\mathfrak{b})(s,X_s,\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ must be 0. We deduce that $(\int_0^t \psi_s b(s,X_s,\mathrm{d}s) = \int_0^t \mathfrak{b}(s,X_s,\mathrm{d}s))_{0\leqslant t\leqslant T_0}$.

We finally discuss the integral $(\int_0^t \psi_s M(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T}$. We let

$$\begin{aligned}\hat{M}(t,t+h) &= X_{t+h} - X_t - \big(B_{t+h} - B_t\big) - \mathbb{E}\big[X_{t+h} - X_t|\mathcal{F}_t\big] \\ &= X_{t+h} - X_t - \big(B_{t+h} - B_t\big) - \mathbb{E}\big[X_{t+h} - X_t - \big(B_{t+h} - B_t\big)|\mathcal{F}_t\big].\end{aligned}$$

By (41), $\mathbb{E}[|\hat{M}(t,t+h)|^q|]^{1/q} \leqslant C_q' h^{(1+\beta)/2}$ for some $C_{q'} \geqslant 0$, which reads as a super-diffusive bound for the pseudo-increments of $\hat{M}$. It is then well-checked that $(\hat{M}(t,t+h))_{0\leqslant t\leqslant t+h\leqslant T_0}$ fulfills all the requirements in (50). Therefore, the integral $(\int_0^t \psi_s \hat{M}(s,s+\mathrm{d}s))_{0\leqslant t\leqslant T_0}$ makes sense. By Subsection 4.5, it is a martingale but by the super-diffusive bound of the pseudo-increments it must be the null process. Put it differently, only the Brownian part really matters in M and we can justify (49) thanks to the equality

$$\int_0^t \psi_s \,\mathrm{d}X_s = \int_0^t \psi_s \,\mathrm{d}B_s + \int_0^t \psi_s b(s,X_s,\mathrm{d}s).$$

5. Construction of the integral of Z w.r.t. Y. Examples.

As a final discussion, we address the existence of a rough path structure $(W_t^T, \mathscr{W}_t^T)_{0\leqslant t\leqslant T}$ for the pair $W_t^T = (Y_t, Z_t^T)$, for T running in some interval $[0,T_0]$, $T_0 > 0$. In this framework, we emphasize that the only challenge is to define the 'cross-integral' $\mathcal{I}_{t,T}(x,x') := \int_x^{x'} (Z_t^T(y) - Z_t^T(x))\,\mathrm{d}Y_t(y)$. Indeed, as we are seeking a geometric rough structure, it makes sense to let (imitating the integration by parts) $\int_x^{x'} (Y_t(y) - Y_t(x))\,\mathrm{d}Y_t(y) := (1/2)(Y_t(x') - Y_t(x))^2$, with a similar identity with Y_t replaced by Z_t^T, and

$$\int_x^{x'} (Y_t(y) - Y_t(x))\,\mathrm{d}Z_t^T(y) := (Y_t(x') - Y_t(x))(Z_t^T(x') - Z_t^T(x)) - \mathcal{I}_{t,T}(x,x'). \tag{60}$$

5.1. **Overview of the results.** We are given $(Y_t(x))_{0\leqslant t\leqslant T_0,x\in\mathbb{R}}$ satisfying for some $\chi,\kappa>0$:

$$\kappa_{\alpha,\chi}((Y_t)_{0\leqslant t\leqslant T_0}):=\sup_{r\geqslant 1,0\leqslant t\leqslant T}\bigl(\|Y_t\|_\alpha^{[-r,r]}/r^\chi\bigr)\leqslant\kappa<\infty,$$

together with $\sup_{0\leqslant t\leqslant T_0}|Y_t(0)|\leqslant\kappa$ (since only the variations of $(Y_t(x))_{0\leqslant t\leqslant T_0,x\in\mathbb{R}}$ matter in (1), we could even assume that $Y_t(0)=0$, for $t\in[0,T_0]$). With the same notation as in (9), we then have $\kappa_{\alpha,\chi}((Y_t,\mathscr{Y}_t)_{0\leqslant t\leqslant T_0})\leqslant C\kappa$, for some universal constant C, where $\mathscr{Y}_t$ denotes the (geometric) iterated integral of Y_t as just defined right above. Recalling the definition (11) of Z_t^T, it can be proved that, for any $T\in[0,T_0]$, $\sup_{0\leqslant t\leqslant T}|Z_t^T(0)|\leqslant C\kappa$ and $\kappa_{\alpha,\chi}((Z_t^T)_{0\leqslant t\leqslant T})\leqslant C\kappa$, the constant C now depending upon T_0 but not on T, thus proving that $\kappa_{\alpha,\chi}((Z_t^T,\mathscr{Z}_t^T)_{0\leqslant t\leqslant T})\leqslant C\kappa$. The proof of $\sup_{0\leqslant t\leqslant T}|Z_t^T(0)|\leqslant C\kappa$ is quite straightforward: Taking benefit of the Hölder regularity of $(Y_t(x))_{0\leqslant t\leqslant T_0,x\in\mathbb{R}}$ in x, the singularity of the second-order derivative of the heat kernel appearing in the definition of Z^T can be integrated. More generally, it can be proved, using the same strategy, that, for any $r\geqslant 1$, $\sup_{0\leqslant t\leqslant T}\sup_{r\geqslant 1}\sup_{x\in[-r,r]}|Z_t^T(x)|\leqslant C\kappa r^\chi$. The proof of $\kappa_{\alpha,\chi}((Z_t^T)_{0\leqslant t\leqslant T})\leqslant C\kappa$ is slightly more subtle. The idea is to go back to (29), with $v\equiv 1$ therein, recalling that the analysis is split into two parts: $|x'-x|^2\leqslant T-t$ and $T-t<|x'-x|^2$, the first case only being challenging. It is then quite straightforward to check that $|\mathcal{I}_1(x,x')(\xi)|\leqslant C\kappa r^\chi|x'-x|^\alpha$, for $x,x',\xi\in[-r,r]$ with $r\geqslant 1$. Moreover, following (31) with $\beta=1$, we also have $\mathcal{I}_2(x,x')\leqslant C\kappa r^\chi\int_x^{x'}\int_{|x-x'|^2}s^{-(3-\alpha)/2}\,\mathrm{d}s\leqslant C\kappa r^\chi|x'-x|^\alpha$, for $x,x'\in[-r,r]$, which completes the proof.

The point is thus to prove that the cross-integral is well-defined and satisfies

$$\bigl|\mathcal{I}_{t,T}(x,x')\bigr|\leqslant C\bigl(1+|x|^{2\chi}+|x'|^{2\chi}\bigr)|x'-x|^{2\alpha},\tag{61}$$

the constant C possibly depending upon T_0 but not on T.

As we already said in Introduction, existence of the cross-integral has been proved within the framework of the KPZ equation by means of general results on rough paths theory applied to Gaussian processes, see [14, Section 7] and [12]. Anyhow, it is a natural question to wonder about the existence for more general classes of 'environments'. In this section, we thus exhibit several sufficient conditions under which (61) is indeed satisfied, the examples we provide being of the following types:

(1) As a first example, we recover the case when the family $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ is constant in time, see Proposition 18.
(2) When the family $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ does depend on time, we prove that the cross-integral is well-defined under some additional time-space regularity. Basically, if the sum of the Hölder exponent in space and twice the Hölder exponent in time is greater than $1-\alpha$, then the cross-integral is well-defined, see Proposition 19. As an application, we deduce that the cross-integral is always well-defined when $\alpha>1/2$, which fits the standard regime for Young's integration.
(3) In Proposition 20, we pay a special attention to the case when $Y_s(x)$ reads, for some $s\geqslant 0$ and $x\in\mathbb{R}$, as a space convolution of $(Y_t(z))_{z\in\mathbb{R}}$ with respect to some heat kernel, for some $t\leqslant s$. Such a situation occurs when the dynamics of Y satisfy a parabolic equation.
(4) The final example we give is of a different nature. We assume that the structure of $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ relies on an additional Brownian sheet so that the cross-integral exists as a stochastic integral. As a basic application, we discuss the case when $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$

is a forward solution to the stochastic heat equation. The case when $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ solves a backward SPDE will be discussed in the next section within the framework of the application to KPZ equation.

It is worth mentioning that the cross-integral is also well-defined when $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ reads as a combination of some of the examples given above.

5.2. **Principle of the analysis.** The construction of the cross-integral may be achieved in two different ways. A first one is to go back to Riemann sums, or possibly compensated Riemann sums as in rough paths theory, and to prove the convergence as the step size of the underlying mesh tends to 0. Another way consists in mollifying the inputs and in proving uniform estimates of the cross-integrals driven by the mollified inputs, the cross-integrals being for sure well-defined in the regular setting. Below, we implement the second strategy as it ensures, for free, the 'geometric' property of the limit cross-integral and as it provides, in some quite interesting cases, explicit expressions of such a limit cross-integral.

We first notice that there is no real difficulty in mollifying the inputs. It is indeed sufficient to mollify first the paths $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ in space and then to plug the mollified version of $(Y_t(x))_{t\geqslant 0,x\in\mathbb{R}}$ in the definition (11) of $Z_t^T(x)$. Below, we consider, as a mollification of $(Y_t(x))_{x\in\mathbb{R}}$, for a given $t\geqslant 0$, the path $Y_t^n(x)=\int_{\mathbb{R}} p_{1/n}(x-y)Y_t(y)\,\mathrm{d}y$. Then, we define $Z_t^{n,T}$ accordingly, by replacing Y_t by Y_t^n in (11).

It is worth mentioning that T is here a given terminal time, which is assumed to live in the compact set $[0,T_0]$. For making the whole machinery work in the previous sections, it is a crucial point to obtain estimates like (61) for the mollified cross-integral that are uniform in $n\geqslant 1$ and but also in t and T, for $0\leqslant t\leqslant T\leqslant T_0$. A first step forward in that direction is to give a sense, for each $s>t$, to the cross integral $I^n_{(t,s)}(x,x'):=\int_x^{x'}\int_{\mathbb{R}}\partial_x^2 p_{s-t}(y-z)Y_s^n(z)\partial_x Y_t^n(y)\,\mathrm{d}y\,\mathrm{d}z$ and to bound it independently of n. In that perspective, a famous result by Young [26] states that, given an exponent $\beta>0$, there exists a universal constant $c>0$ such that, for any two smooth functions f and g on the interval $[x,x']$ satisfying

$$|f(z)-f(y)|\,|g(z')-g(y)|\leqslant C|z-z'|^{1+\beta}, \tag{62}$$

for some constant C and for any $x\leqslant z\leqslant y\leqslant z'\leqslant x'$, it holds

$$\left|\int_x^{x'} f(z)g'(z)\,\mathrm{d}z - f(x)\big(g(x')-g(x)\big)\right|\leqslant cC|x'-x|^{1+\beta}. \tag{63}$$

Letting $f(y)=\int_{\mathbb{R}}\partial_x^2 p_{s-t}(y-z)Y_s^n(z)\,\mathrm{d}z$ and $g(y)=Y_t^n(y)$, we notice that f is α'-Hölder continuous on $[x,x']$, for $0<\alpha'\leqslant 1$, with a Hölder constant of order $(s-t)^{-(1+\alpha'/2)+\alpha/2}$ (independently of n). Equation (62) then holds with $1+\beta=\alpha'+\alpha$ and C of order $(s-t)^{-(1+\alpha'/2)+\alpha/2}$ (independently of n). Asking $\beta>0$ requires $\alpha'+\alpha>1$ and asking the Hölder constant to be integrable (uniformly in n) requires $\alpha'<\alpha$. In the end, for defining the cross-integral, Young's theory only applies if $\alpha>1/2$.

Below, we go back to Young's framework in a more detailed, but generally speaking, the objective is to go further into the analysis and to discuss several cases when the cross-integral exists even if $\alpha\leqslant 1/2$ (α being larger than $1/3$).

5.3. **Time homogeneous environment.** The first step is to discuss the case when Y is constant in time or, equivalently, to focus, in the time-dependent framework, on the mollified

cross-integral $\int_x^{x'} \mathcal{Z}_t^{n,T}(y)\partial_x Y_t^n(y)\,\mathrm{d}y$, with

$$\mathcal{Z}_t^{n,T}(y) := \int_t^T \int_{\mathbb{R}} \partial_s p_{s-t}(y-z) Y_t^n(z)\,\mathrm{d}z\,\mathrm{d}s = \int_{\mathbb{R}} p_{T-t}(y-z) Y_t^n(z)\,\mathrm{d}z - Y_t^n(y).$$

Here $\mathcal{Z}_t^{n,T}$ reads as a modified version of $Z_t^{n,T}$, in which Y_s^n has been replaced by Y_t^n (see (11)). In comparison with (11), the second order derivative $\partial_x^2 p_{s-t}$ has been also replaced by $\partial_s p_{s-t}$, by taking benefit (up to a factor 2) of the heat equation satisfied by p_{s-t}.

In such a case, the cross-integral $\int_x^{x'} (\mathcal{Z}_t^{n,T}(y) - \mathcal{Z}_t^{n,T}(x))\,\mathrm{d}Y_t^n(y)$ can be expressed explicitly, without any reference to the derivative of Y_t^n:

$$\begin{aligned}
&\int_x^{x'} \big(\mathcal{Z}_t^{n,T}(y) - \mathcal{Z}_t^{n,T}(x)\big)\,\mathrm{d}Y_t^n(y)\\
&= \int_x^{x'} \int_{\mathbb{R}} \big(p_{T-t}(y-z) - p_{T-t}(x-z)\big) Y_t^n(z)\partial_x Y_t^n(y)\,\mathrm{d}z\,\mathrm{d}y\\
&\quad - \int_x^{x'} \big(Y_t^n(y) - Y_t^n(x)\big)\partial_x Y_t^n(y)\,\mathrm{d}y\\
&= Y_t^n(x')\big(\mathcal{Z}_t^{n,T}(x') - \mathcal{Z}^{n,T}(x) + Y_t^n(x') - Y_t^n(x)\big) - \frac{1}{2}\big[\big(Y_t^n(x') - Y_t^n(x)\big)^2\big]\\
&\quad - \int_x^{x'} \int_{\mathbb{R}} \partial_x p_{T-t}(y-z) Y_t^n(z) Y_t^n(y)\,\mathrm{d}z\,\mathrm{d}y,
\end{aligned} \tag{64}$$

the passage from the second to the third line following from an integration by parts. It is quite straightforward to see that the cross-integrals converge as n tends to the infinity, the limit cross-integral satisfying a similar formula. A crucial point is that it satisfies (61), uniformly in $t \leqslant T$ in $[0, T_0]$:

Proposition 18. *Under the assumptions and notations specified in Subsection 5.1, let* $\mathcal{Z}_t^T(y) = \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(y-z) Y_t(z)\,\mathrm{d}z\,\mathrm{d}s$. *Then, the integral* $\int_x^{x'} (\mathcal{Z}_t^T(y) - \mathcal{Z}_t^T(x))\,\mathrm{d}Y_t(y)$ *may be defined as a 'geometric integral' obtained by replacing* $(Y^n, \mathcal{Z}^{n,T})$ *by* $(Y, \mathcal{Z}^T)$ *in* (64). *It satisfies* (61), *the constant* C *therein being uniform with respect to* $t \leqslant T$ *in* $[0, T_0]$.

Proof. By the mollification argument, it is sufficient to prove that the bound (61) holds for Y smooth, provided the resulting constant C in (61) only depends on κ and T_0. Replacing $Y_t^n(y)$ by $Y_t(y) - Y_t(x)$ in (64) (which doesn't change the definition of $\mathcal{Z}_t^{n,T}$) and subtracting $\mathcal{Z}_t^T(x)[Y_t(x') - Y_t(x)]$, we get

$$\begin{aligned}
\mathcal{I}_{t,T}(x,x') &= \big(Y_t(x') - Y_t(x)\big)\big(\mathcal{Z}_t^T(x') - \mathcal{Z}_t^T(x) + Y_t(x') - Y_t(x)\big) - \frac{1}{2}\big(Y_t(x') - Y_t(x)\big)^2\\
&\quad - \int_x^{x'} \int_{\mathbb{R}} \partial_x p_{T-t}(y-z)\big(Y_t(z) - Y_t(x)\big)\big(Y_t(y) - Y_t(x)\big)\,\mathrm{d}z\,\mathrm{d}y.
\end{aligned}$$

The terms in the first line are easily tackled. Indeed, the term $\mathcal{Z}_t^T$ has the same regularity in space as Z_t^T (the proof is the same, replacing Y_s by Y_t). The point is thus to investigate the last term in the definition of $\mathcal{I}_{t,T}(x,x')$. We put things in a more general framework that will be reused in the sequel. Letting $h = T - t$ in the definition of $\mathcal{I}_{t,T}(x,x')$, we

replace $\partial_x p_h(y-z)$ by another antisymmetric function $G_h(y,z)$ (in the sense that $G_h(z,y) = -G_h(y,z)$) satisfying, for any $\beta \geqslant 0$,

$$\int_{\mathbb{R}} |y-z|^\beta \big|G_h(y,z)\big| \,\mathrm{d}z \leqslant C_\beta h^{\beta/2-1/2}, \tag{65}$$

and, for any $a \in \mathbb{R}$,

$$\int_{-\infty}^a \biggl(\int_a^{+\infty} |y-z|^\beta \big|G_h(y,z)\big| \,\mathrm{d}z \biggr) \,\mathrm{d}y + \int_a^\infty \biggl(\int_{-\infty}^a |y-z|^\beta \big|G_h(y,z)\big| \,\mathrm{d}z \biggr) \,\mathrm{d}y \leqslant C_\beta h^{\beta/2}, \tag{66}$$

the constant C_β possibly depending on T_0. It is standard to check that $G_h(y,z) = \partial_x p_h(y-z)$ satisfies (65). The verification of (66) is a bit more involved. Since $\partial_x p_h(y-z) \geqslant 0$ for $y \geqslant z$, we have (the value of C_β varying from line to line)

$$\begin{aligned}
\int_a^\infty \biggl(\int_{-\infty}^a |y-z|^\beta \big|\partial_x p_h(y-z)\big| \,\mathrm{d}z \biggr) \,\mathrm{d}y &= \int_a^\infty \biggl(\int_{-\infty}^a h^{-1} \bigl(y-z\bigr)^{1+\beta} p_1\bigl(\frac{y-z}{h^{1/2}}\bigr) \,\mathrm{d}z \biggr) \,\mathrm{d}y \\
&\leqslant C_\beta h^{\beta/2} \int_a^\infty \biggl(\int_{-\infty}^a h^{-1} \bigl(y-z\bigr) p_2\bigl(\frac{y-z}{h^{1/2}}\bigr) \,\mathrm{d}z \biggr) \,\mathrm{d}y \\
&\leqslant C_\beta h^{\beta/2} \int_a^\infty h^{-1/2} p_2\bigl(\frac{y-a}{h^{1/2}}\bigr) \,\mathrm{d}y.
\end{aligned}$$

We then consider:

$$\mathcal{I}_h(x,x') := \int_x^{x'} \int_{\mathbb{R}} G_h(y,z) \bigl[Y_t(z) - Y_t(x)\bigr] \bigl[Y_t(y) - Y_t(x)\bigr] \,\mathrm{d}z \,\mathrm{d}y, \tag{67}$$

with $-r \leqslant x < x' \leqslant r$, for some $r \geqslant 1$. Splitting $Y_t(z) - Y_t(x)$ into $Y_t(z) - Y_t(y)$ plus $Y_t(y) - Y_t(x)$, we deduce from (65) that $|\mathcal{I}_h(x,x')| \leqslant C(1+r^{2\chi})[h^{-1/2}|x'-x|^{1+2\alpha} + h^{-(1-\alpha)/2}|x'-x|^{1+\alpha}]$, so that, for $h \geqslant |x'-x|^2$,

$$|\mathcal{I}_h(x,x')| \leqslant C\bigl(1+r^{2\chi}\bigr)|x'-x|^{2\alpha}.$$

In order to handle the case $h \leqslant |x'-x|^2$, we first deduce from the antisymmetry property that

$$\int_x^{x'} \int_x^{x'} G_h(y,z) \bigl[Y_t(z) - Y_t(x)\bigr] \bigl[Y_t(y) - Y_t(x)\bigr] \,\mathrm{d}z \,\mathrm{d}y = 0.$$

Therefore, $\mathcal{I}_h(x,x') = \mathcal{I}_h^1(x,x') + \mathcal{I}_h^2(x,x')$, with

$$\begin{aligned}
\mathcal{I}_h^1(x,x') &:= \int_x^{x'} \int_{x'}^{+\infty} G_h(y,z) \bigl[Y_t(z) - Y_t(x)\bigr] \bigl[Y_t(y) - Y_t(x)\bigr] \,\mathrm{d}z \,\mathrm{d}y, \\
\mathcal{I}_h^2(x,x') &:= \int_x^{x'} \int_{-\infty}^{x} G_h(y,z) \bigl[Y_t(z) - Y_t(x)\bigr] \bigl[Y_t(y) - Y_t(x)\bigr] \,\mathrm{d}z \,\mathrm{d}y.
\end{aligned}$$

We start with $\mathcal{I}_h^1(x,x')$. We write $\mathcal{I}_h^1(x,x') = \mathcal{I}_h^{1,1}(x,x') + \mathcal{I}_h^{1,2}(x,x')$, with

$$\begin{aligned}
\mathcal{I}_h^{1,1}(x,x') &:= \int_x^{x'} \int_{x'}^{+\infty} G_h(y,z) \bigl[Y_t(z) - Y_t(y)\bigr] \bigl[Y_t(y) - Y_t(x)\bigr] \,\mathrm{d}z \,\mathrm{d}y, \\
\mathcal{I}_h^{1,2}(x,x') &:= \int_x^{x'} \int_{x'}^{+\infty} G_h(y,z) \bigl[Y_t(y) - Y_t(x)\bigr]^2 \,\mathrm{d}z \,\mathrm{d}y.
\end{aligned}$$

By (66), we have $|\mathcal{I}_h^{1,1}(x,x')| \leqslant C(1+r^{2\chi})h^{\alpha/2}|x-x'|^\alpha$. Similarly, $|\mathcal{I}_h^{1,2}(x,x')| \leqslant C(1+r^{2\chi})|x-x'|^{2\alpha}$ Therefore,

$$\big|\mathcal{I}_h^1(x,x')\big| \leqslant C\big(1+r^{2\chi}\big)\big(h^{\alpha/2}|x-x'|^\alpha + |x-x'|^{2\alpha}\big).$$

Notice that $\mathcal{I}_h^2(x,x')$ can be bounded in a similar way. Therefore,

$$\big|\mathcal{I}_h(x,x')\big| \leqslant C\big(1+r^{2\chi}\big)\big(h^{\alpha/2}|x-x'|^\alpha + |x-x'|^{2\alpha}\big),$$

which is less than $C(1+r^{2\chi})|x-x'|^{2\alpha}$ if $h \leqslant |x-x'|^2$.

□

5.4. **Time space regularity.** From the practical point of view, the previous paragraph is of real interest as its scope goes beyond the time homogeneous framework. Indeed, it says that, in the analysis of the mollified cross-integrals –see Subsection 5.2 –, one can limit ourselves to the investigation of $\int_x^{x'}(\mathcal{Z}_t^{n,T}(y)-\mathcal{Z}_t^{n,T}(x))\partial_x Y_t^n(y)\,\mathrm{d}y$, with a new of definition of $\mathcal{Z}_t^{n,T}(y)$ (which, we feel, yields no confusion with the one introduced in the previous subsection):

$$\mathcal{Z}_t^{n,T}(y) := \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(y-z)\big(Y_s^n(z)-Y_t^n(z)\big)\,\mathrm{d}z\,\mathrm{d}s. \tag{68}$$

Put it differently, one can distinguish the Y^n appearing in the cross-integral from the Y^n appearing in the definition of $\mathcal{Z}_t^{n,T}$. Below, we investigate several cases of a more general kind:

$$\mathcal{Z}_t^{n,T}(y) = \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(y-z)\mathcal{Y}_{t,s}^n(z)\,\mathrm{d}z\,\mathrm{d}s, \tag{69}$$

that covers (68), $(\mathcal{Y}_{t,s}^n(y))_{0\leqslant t\leqslant s\leqslant T_0, y\in\mathbb{R}}$ denoting a family of paths indexed by two-dimensional time indices.

5.4.1. *Young's theory.* Making use of Young's theory, we are then able to prove:

Proposition 19. *Let the assumptions and notations specified in Subsection 5.1 be in force. Consider also a time-space family $(\mathcal{Y}_{t,s}(y))_{0\leqslant t\leqslant s\leqslant T_0, y\in\mathbb{R}}$ such that, for some $\mu,\mu'\geqslant 0$ satisfying $2\mu'+\mu>1-\alpha$ and for some constant $\kappa'>0$, it holds*

$$\forall s\in[t,T_0],\ \forall y,z\in\mathbb{R},\quad |\mathcal{Y}_{t,s}(z)-\mathcal{Y}_{t,s}(y)| \leqslant \kappa'\big(1+|y|^\chi+|z|^\chi\big)|s-t|^{\mu'}|y-z|^\mu. \tag{70}$$

For $0\leqslant t\leqslant T\leqslant T_0$, let

$$\mathcal{Z}_t^T(y) := \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(y-z)\mathcal{Y}_{t,s}(z)\,\mathrm{d}z\,\mathrm{d}s, \tag{71}$$

and define, on the same model as in Subsection 5.2, a mollified version $\mathcal{Y}_{t,s}^n$ of $\mathcal{Y}_{t,s}$ by convolution with the heat kernel and then consider $\mathcal{Z}_t^{n,T}$ as in (69)*. Then, for any $\alpha'\in(1-\alpha,\mu+2\mu')$, there exists a constant $C'=C'(\alpha',\kappa,\kappa',T_0)$ such that, for any $0\leqslant t\leqslant T\leqslant T_0$, the 'geometric integral' of $\mathcal{Z}_t^T$ with respect to Y_t makes sense and satisfies* (61) *with respect to C' and 2α replaced by $\alpha+\alpha'$.*

The reader might worry about the fact that (61) holds with respect to a larger exponent than 2α. The resulting effect can be read in Theorem 1, in which the regularity of $\mathscr{W}$ explicitly appears. As a consequence, it affects Lemma 2 as it generates a new term in the definition of $\mathcal{D}$, which writes (with the same notation as therein) $|z|^{2(\alpha+\alpha')+\beta}\rho^{2\chi+\beta/2}$. This new definition of $\mathcal{D}$ must be injected in the proof of Lemma 11. Basically, a new term must

be added to the definition of $\mathcal{D}'$ in (20): It is of the form $s^{\alpha+\alpha'+\beta/2-1}\rho^{2\chi+\beta/2}$. The important point is that the power of ρ is not changed in comparison with the original case when $\alpha' = 0$. Thus, the final result remains true.

A typical example of application is $\mathcal{Y}_{t,s}(y) = Y_s(y) - Y_t(y)$. In the specific case when $Y_t(y)$ may be expanded as $Y_t(y) = f_t Y(y)$, f being more than $(1/2-\alpha)$-Hölder continuous, (70) holds with $\mu = \alpha$ and $\mu' > 1/2 - \alpha$. Another example of application is $\mathcal{Y}_{t,s}(y) = Y_s(y)$. In that framework, (70) is satisfied with $\mu > 1/2$ and $\mu' = 0$ if $\alpha > 1/2$, in which case the cross-integral fits the Young theory discussed in Subsection 5.2.

Proof. As in the proof of Proposition 18, we can assume that the mappings Y_t and $(\mathcal{Y}_{t,s})_{t\leqslant s\leqslant T}$ are smooth in space. The proof is then an application of Young estimates (62) and (63), with $f(y) = \int_{\mathbb{R}} \partial_x^2 p_{s-t}(y-z)\mathcal{Y}_{t,s}(z)\,\mathrm{d}z$ and $g(y) = Y_t(y)$. We notice that f is α'-Hölder continuous on $[x, x']$, for $0 < \alpha' \leqslant 1$, with a Hölder constant of the form $C'(1 + |x|^\chi + |x'|^\chi)(s-t)^{-(1+\alpha'/2)+\mu/2+\mu'}$ (C' depending upon α'). Eq. (62) then holds with $1 + \beta = \alpha' + \alpha$ and C of the form $C'(s-t)^{-(1+\alpha'/2)+\mu/2+\mu'}$. Asking $\beta > 0$ requires $\alpha' + \alpha > 1$ and asking the Hölder constant to be integrable requires $\alpha' < \mu + 2\mu'$. Young's theory applies if $\mu + 2\mu' + \alpha > 1$, in which case (61) holds with 2α replaced by $\alpha + \alpha' > 1$.

□

5.4.2. *Space convolution.* As another important example for practical applications, we consider the case when $\mathcal{Y}_{t,s}$ is given as a convolution (in space) of the increment $Y_s - Y_t$:

Proposition 20. *Let the assumptions and notations specified in Subsection 5.1 be in force. Consider also a family of kernels $(q_{t,s}(z,u))_{0\leqslant t<s\leqslant T_0, z,u\in\mathbb{R}}$ (that is a family of non-negative functions with mass 1 in u) satisfying:*

$$(72)\quad \Big|q_{t,s}(z,u) - p_{(s-t)\sigma_t^2(z)}(z-u)\Big| \leqslant C(s-t)^\mu p_{C^2(s-t)}(z-u), \quad 0 \leqslant t < s \leqslant T_0, \quad z,u\in\mathbb{R},$$

with $\mu > 1/2 - \alpha$ and $C \geqslant 1$, $\sigma_t(\cdot)$ being a C-Lipschitz continuous function with values in $[1/C, C]$. Letting $\mathcal{Y}_{t,s}(z) = \int_{\mathbb{R}} q_{t,s}(z,u)(Y_t(u) - Y_t(z))\,\mathrm{d}u$ and defining $\mathcal{Y}^n_{t,s}(z)$ by mollifying $\mathcal{Y}_{t,s}$ with a kernel of variance $1/n$, consider $\mathcal{Z}^T_t(y)$ given by (71) and $\mathcal{Z}^{n,T}_t(y)$ given by (69). Then, there exist a constant $C' = C'(\mu,\kappa,C,T_0)$ and an exponent $\epsilon = \epsilon(\alpha,\mu) > \max(2\alpha - 1, 0)$ such that, for any $0 \leqslant t \leqslant T \leqslant T_0$, the 'geometric integral' of $\mathcal{Z}^T_t$ with respect to Y_t makes sense and satisfies (61) with $C'r^{2\chi}|x'-x|^{2\alpha}$ replaced by $C'r^{2\chi}(|x'-x|^{2\alpha} + |x'-x|^{1+\epsilon})$.

As emphasized right after Proposition 19, the reader must not worry about the additional $|x'-x|^{1+\epsilon}$.

Example 21. *An example for q is the transition density of a second-order operator $L_t = b_t(\cdot)\partial_x + (1/2)\sigma_t^2(\cdot)\partial_x^2$, where b is a bounded function and σ is, in addition to the assumption specified in the statement of Proposition 20, $1/2$-Hölder continuous in time, uniformly in space. The proof of (72) follows from the 'parametrix' method in* [11, Chapter 1].

Proof of Proposition 20. First step. The first step of the proof consists in showing that $\mathcal{Y}^n_{s,t}$ can be chosen as the integral with respect to q of the mollified version Y^n_t of Y_t. With the definition given in the statement (which is the right one for defining a 'geometric integral'),

$$\mathcal{Y}^n_{t,s}(z) = \int_{\mathbb{R}} p_{1/n}(z-v)\mathcal{Y}^n_{t,s}(v)\,\mathrm{d}v = \int_{\mathbb{R}} q^n_{t,s}(z,u)\big(Y_t(u) - Y^n_t(z)\big)\,\mathrm{d}u,$$

with $q^n_{t,s}(z,u) = \int_{\mathbb{R}} p_{1/n}(z-v) q_{t,s}(v,u)\,\mathrm{d}z$. The point is then to estimate $q^n_{t,s}(z,u)$ by taking benefit of (72) and of a Gaussian convolution. The argument is based on the fact that, for $\sigma \in [1/C, C]$ and $0 < s-t \leqslant T_0$,

$$\big|\partial_\sigma\big[p_{\sigma^2(s-t)}(z-u)\big]\big| = \big|\partial_\sigma\big[\sigma^{-1} p_{s-t}\big(\frac{z-u}{\sigma}\big)\big]\big| \leqslant C' p_{s-t}\big(\frac{z-u}{C'}\big), \tag{73}$$

for some constant $C' > 0$. This says that, in (72), we can replace $\sigma_t(z)$ by $\sigma_t(u)$ up to a modification of the value of the constant C in the right-hand side. Indeed, the distance $|\sigma_t(u) - \sigma_t(z)|$ is less than $C|u-z|$, which is of order $(s-t)^{1/2}$ when multiplied by the heat kernel of variance $s-t$.

Recalling the very basic formula $\int_{\mathbb{R}} p_{s_1}(y-z) p_{s_2}(z)\,\mathrm{d}z = p_{s_1+s_2}(y)$, for $y \in \mathbb{R}$ and $s_1, s_2 > 0$, we then deduce from the new version of (72) that

$$\begin{aligned}\big|q^n_{t,s}(z,u) - p_{1/n+(s-t)\sigma_t^2(u)}(z-u)\big| &= \Big|\int_{\mathbb{R}} p_{1/n}(z-v)\big[q_{t,s}(v,u) - p_{(s-t)\sigma_t^2(u)}(v-u)\big]\,\mathrm{d}v\Big| \\ &\leqslant C(s-t)^\mu p_{1/n+C^2(s-t)}(z-u).\end{aligned}$$

Following (73) and modifying the constant C if necessary, the same holds true with $\sigma_t^2(u)$ replaced by $\sigma_t^2(z)$, that is

$$\big|q^n_{t,s}(z,u) - p_{1/n+(s-t)\sigma_t^2(z)}(z-u)\big| \leqslant C(s-t)^\mu p_{1/n+C^2(s-t)}(z-u).$$

We then notice that

$$\begin{aligned}\int_{\mathbb{R}} p_{1/n+(s-t)\sigma_t^2(z)}(z-u)\big(Y_t(u) - Y_t^n(z)\big)\,\mathrm{d}u &= \int_{\mathbb{R}} p_{(s-t)\sigma_t^2(z)}(z-v)\big(Y_t^n(v) - Y_t^n(z)\big)\,\mathrm{d}v, \\ \int_{\mathbb{R}} p_{1/n+C^2(s-t)}(z-u)\big(Y_t(u) - Y_t^n(z)\big)\,\mathrm{d}u &= \int_{\mathbb{R}} p_{C^2(s-t)}(z-v)\big(Y_t^n(v) - Y_t^n(z)\big)\,\mathrm{d}v,\end{aligned}$$

from which we deduce that

$$\mathcal{Y}^n_{t,s}(z) = \int_{\mathbb{R}} p_{(s-t)\sigma_t^2(z)}(z-u)\big(Y_t^n(u) - Y_t^n(z)\big)\,\mathrm{d}u + \epsilon^n_{s,t}(z), \tag{74}$$

where $\epsilon^n_{s,t}(z)$ denotes a remainder, bounded by $C'(1+|z|^\chi)(s-t)^{\alpha/2+\mu}$, the constant C' being independent of n. The integral with respect to the remainder can be estimated by means of Proposition 19 since $\alpha/2 + \mu > (1-\alpha)/2$. This says that we can forget the remainder $\epsilon^n_{s,t}$ and do as if it was 0.

Below, we thus focus on the cross-integral driven by the first term only in the right-hand side in (74). This permits to do the same as in the previous proofs: We can directly assume that Y_t is smooth in space and forget the superscript n in the notations. Regardless the restriction $\epsilon^n_{s,t} \equiv 0$, the cross-integral is still denoted by $\int_x^{x'} (\mathcal{Z}_t^T(y) - \mathcal{Z}_t^T(x))\,\mathrm{d}Y_t(y)$.

Second step. The point is to prove (61) for $-r \leqslant x < x' \leqslant r$, with $r \geqslant 1$. By integration by parts (see the introduction of the section), it is equivalent to focus on $\int_x^{x'} (Y_t(y) - Y_t(x))\,\mathrm{d}\mathcal{Z}_t^T(y)$, which reads

$$\int_t^T \int_x^{x'} \int_{\mathbb{R}} \partial_x^3 p_{s-t}(y-z)\Big(\int_{\mathbb{R}} p_{(s-t)\sigma_t^2(z)}(z-u)\big(Y_t(u) - Y_t(z)\big)\big(Y_t(y) - Y_t(x)\big)\,\mathrm{d}u\Big)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s.$$

By (73), we can replace $\sigma_t(z)$ by $\sigma_t(y)$. Indeed, the function σ_t being Lipschitz in space, the difference between the two terms can be bounded by (for a new value of C')

$$C'\big(1+r^{2\chi}\big)\int_t^T\int_x^{x'}(s-t)^{-3/2+1/2+\alpha/4}|y-x|^\alpha\,\mathrm{d}y\,\mathrm{d}s,$$

which is less than $C'(1+r^{2\chi})(T-t)^{\alpha/4}|x-x'|^{1+\alpha}$. Actually, it is even sufficient to focus on

$$\int_t^T\int_x^{x'}\int_{\mathbb{R}}\partial_x^3 p_{s-t}(y-z)\int_{\mathbb{R}}p_{(s-t)\sigma_t^2(y)}(z-u)Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s, \tag{75}$$

as the remainder, which reads $\int_t^T\int_x^{x'}\int_{\mathbb{R}}\partial_x^3 p_{s-t}(y-z)Y_t(z)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}s$, has been already tackled within the framework of time-homogeneous environments.

Third step. We denote by $\mathcal{J}_{t,T}(x,x')$ the term in (75). The good point is that, by convolution in z, the expression can be reduced to

$$\int_t^T\int_x^{x'}\int_{\mathbb{R}}\frac{d^3}{dw^3}_{|w=y}p_{(s-t)[1+\sigma_t^2(y)]}(w-u)Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}s\,\mathrm{d}y\,\mathrm{d}u.$$

Noting that $\partial_x^2 p_{(s-t)}(\cdot)=2\partial_s[p_{s-t}]$, we get that $\mathcal{J}_{t,T}(x,x')$ is equal to

$$2\int_t^T\int_x^{x'}\int_{\mathbb{R}}[1+\sigma_t^2(y)]^{-2}\partial_s\partial_x p_{s-t}\big(\frac{y-u}{[1+\sigma_t^2(y)]^{1/2}}\big)Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y\,\mathrm{d}s.$$

which, by integration by parts in u, also matches

$$2\int_t^T\int_x^{x'}\int_{\mathbb{R}}[1+\sigma_t^2(y)]^{-3/2}\partial_s p_{s-t}\big(\frac{y-u}{[1+\sigma_t^2(y)]^{1/2}}\big)\partial_x Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y\,\mathrm{d}s.$$

Integrating in time, we finally get

$$\begin{aligned}\mathcal{J}_{t,T}(x,x')&=2\int_x^{x'}\int_{\mathbb{R}}[1+\sigma_t^2(y)]^{-3/2}p_{T-t}\big(\frac{y-u}{[1+\sigma_t^2(y)]^{1/2}}\big)\partial_x Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y\\&\quad-2\int_x^{x'}[1+\sigma_t^2(y)]^{-1}\partial_x Y_t(y)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}y.\end{aligned}$$

The second term in the right-hand side writes

$$-\int_x^{x'}[1+\sigma_t^2(y)]^{-1}\frac{d}{dy}\big[\big(Y_t(y)-Y_t(x)\big)^2\big]\,\mathrm{d}y,$$

and easily tackled by integration by parts, taking benefit of the Lipschitz property of σ. It is less than $Cr^{2\chi}(|x'-x|^{2\alpha}+|x'-x|^{1+2\alpha})$.

Fourth step. It thus remains to discuss

$$2\int_x^{x'}\int_{\mathbb{R}}[1+\sigma_t^2(y)]^{-3/2}p_{T-t}\big(\frac{y-u}{[1+\sigma_t^2(y)]^{1/2}}\big)\partial_x Y_t(u)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y.$$

Performing an integration by parts in u and using the Lipschitz property of σ, it is equal to

$$2\int_x^{x'}\int_{\mathbb{R}}[1+\sigma_t^2(y)]^{-2}\partial_x p_{T-t}\big(\frac{y-u}{[1+\sigma_t^2(y)]^{1/2}}\big)\big(Y_t(u)-Y_t(x)\big)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y.$$

Using the analogue (73) for $\partial_\sigma \partial_x p_{s-t}$ (which consists in adding a constant $(s-t)^{-1/2}$ in front of the right-hand side), we can replace $\sigma_t^2(y)$ by $\sigma_t(y)\sigma_t(u)$. The resulting difference is indeed less than $Cr^{2\chi}(|x'-x|^{1+\alpha} + |x'-x|^{1+2\alpha})$. We are thus left with

$$2\int_x^{x'} \int_{\mathbb{R}} [1+\sigma_t(y)\sigma_t(u)]^{-2} \partial_x p_{T-t}\Big(\frac{y-u}{[1+\sigma_t(y)\sigma_t(u)]^{1/2}}\Big)\big(Y_t(u)-Y_t(x)\big)\big(Y_t(y)-Y_t(x)\big)\,\mathrm{d}u\,\mathrm{d}y.$$

It can be put in the same form as in (67), with

$$G_h(y,z) = [1+\sigma_t(y)\sigma_t(z)]^{-2}\partial_x p_h\Big(\frac{y-z}{[1+\sigma_t(y)\sigma_t(z)]^{1/2}}\Big).$$

It is thus sufficient to check (65) and (66), which is not very difficult since, for $y \geqslant z$ (and similarly for $z \leqslant y$), $0 \leqslant G_h(y,z) \leqslant C'\partial_x p_h[(y-z)/C']$, for some constant C'.

In order to complete the proof, we notice that all the exhibited bounds are of the form $Cr^\chi|x'-x|^\gamma$, for γ varying between $\gamma_{\min} = 2\alpha$ and some $\gamma_{\max} > 1$. By standard Young's inequality, we can bound any $|x'-x|^\gamma$ by $C(|x'-x|^{\gamma_{\min}} + |x'-x|^{\gamma_{\max}})$. □

5.5. **Decorrelation of the increments.** Another strategy is to assume that $\mathcal{Y}$ is random, with suitable decorrelation properties, so that the cross-integral can be constructed as a stochastic integral. Denoting by $(\Xi, \mathcal{G}, \mathbf{P})$ the corresponding probability space (which models some random 'environment' and which is given apart from the space $(\Omega, \mathcal{A}, \mathbb{P})$ used to build the solution to (1)), we claim:

Proposition 22. *Let $(\zeta(t,x))_{0\leqslant t\leqslant T_0, x\in\mathbb{R}}$ be a Brownian sheet with respect to some filtration $(\mathcal{G}_t)_{0\leqslant t\leqslant T_0}$ on $(\Xi,\mathcal{G},\mathbf{P})$, such that, for any $t\in[0,T_0]$, the mapping Y_t (seen as a random function) is $\mathcal{G}_t$-measurable. Assume also that, for any $T\in[0,T_0]$, there exists a $\mathcal{G}_T$-measurable random variable κ_T, with $\sup_{0\leqslant T\leqslant T_0}\mathbf{E}[|\kappa_T|^p] < +\infty$ for all $p\geqslant 1$, such that, for all $r\geqslant 1$,*

$$\forall t\in[0,T],\ \forall x,x'\in[-r,r],\quad |Y_t(x)-Y_t(x')| \leqslant \kappa_T r^\chi |x-x'|^\alpha,$$

with $\alpha\in(1/3,1/2)$, and

$$\forall t,t'\in[0,T],\ \forall x\in[-r,r],\quad |Y_t(x)-Y_{t'}(x)| \leqslant \kappa_T r^{\chi'}|t'-t|^\eta,$$

with $\chi\in(0,\alpha/2)$ and $\chi',\eta>0$. In addition let $(k_{t,s}(x,y)_{0\leqslant t\leqslant s\leqslant T_0, x,y\in\mathbb{R}}$ be a progressively-measurable process, in the sense that the mapping $\Delta_{T,T_0}\times\mathbb{R}^2 \ni (t,s,x,y)\mapsto k_{t,s}(x,y)\in\mathbb{R}$ is $\mathcal{G}_T\otimes\mathcal{B}(\Delta_{T,T_0}\times\mathbb{R}^2)$ measurable, where $\Delta_{T,T_0} := \{(t,s)\in[0,T]\times[0,T_0] : t\leqslant s\}$. Assume that it satisfies

$$|k_{t,s}(x,y)| \leqslant C p_{s-t}\Big(\frac{y-x}{C}\Big),\quad |\partial_x k_{t,s}(x,y)| \leqslant C(s-t)^{-1/2} p_{s-t}\Big(\frac{y-x}{C}\Big).$$

Letting $\mathcal{Y}_{t,s}(z) = \int_t^s\int_{\mathbb{R}} k_{\rho,s}(z,u)\,\mathrm{d}\zeta(\rho,u)$ and defining $\mathcal{Y}^n_{t,s}(z)$ in a similar way by mollification, consider $\mathcal{Z}_t^T(y)$ given by (71) and $\mathcal{Z}_t^{n,T}(y)$ given by (69). Then, for any $\alpha'\in(0,\alpha)$, there exist an exponent $\chi'=\chi'(\alpha',\alpha,\eta,\chi)$, with $\chi'\to\chi$ as $\alpha'\to\alpha$, and a random variable κ', with finite moments of any order, such that, almost surely, for any $0\leqslant t\leqslant T\leqslant T_0$, the 'geometric integral' of $\mathcal{Z}_t^T$ with respect to Y_t makes sense and satisfies (61) with respect to C replaced by κ', α replaced by α' and χ replaced by χ'.

Proof. First step. We first assume that $\mathcal{Z}_t^T(y)$ and $Y_t(y)$ are regular in the variable y. By (60), it is enough to focus on the cross-integral $\mathcal{J}_{t,T}(x,x') := \int_x^{x'}[Y_t(y)-Y_t(x)]\partial_x\mathcal{Z}_t^T(y)\,\mathrm{d}y$ instead of $\mathcal{I}_{t,T}(x,x') = \int_x^{x'}[\mathcal{Z}_t^T(y)-\mathcal{Z}_t^T(x)]\partial_x Y_t(y)\,\mathrm{d}y$.

By stochastic Fubini's theorem and by a change of variable, we can write $\mathcal{J}_{t,T}(x,x') = \int_t^T \mathcal{H}_{t,s}(x,x')\,\mathrm{d}s$, with

$$\mathcal{H}_{t,s}(x,x') = \int_{\mathbb{R}} \partial_x^3 p_{s-t}(z)\biggl[\int_t^s \int_{\mathbb{R}}\biggl(\int_x^{x'} k_{\rho,s}(y-z,u)\bigl[Y_t(y)-Y_t(x)\bigr]\,\mathrm{d}y\biggr)\,\mathrm{d}\zeta(\rho,u)\biggr]\,\mathrm{d}z. \tag{76}$$

Clearly, we can bound $|\mathcal{J}_{t,T}(x,x')|$ by $\mathcal{K}_t(x,x') := \int_t^{T_0} |\mathcal{H}_{t,s}(x,x')|\,\mathrm{d}s$. The strategy is then to bound the moments of $\mathcal{K}_t(x,x')$ and thus to compute the bracket of the stochastic integral in (76). The bracket reads

$$A_s(z) := \int_t^s \int_{\mathbb{R}} \int_x^{x'} \int_x^{x'} k_{\rho,s}(y-z,u)k_{\rho,s}(y'-z,u)\bigl(Y_t(y)-Y_t(x)\bigr)\bigl(Y_t(y')-Y_t(x)\bigr)\,\mathrm{d}\rho\,\mathrm{d}u\,\mathrm{d}y\,\mathrm{d}y'.$$

Assuming that $-r \leqslant x < x' \leqslant r$, for some $r \geqslant 1$, and recalling that there exists a $\mathcal{G}_t$-measurable random variable κ_t, with $\sup_{0\leqslant t\leqslant T_0} \mathbf{E}[|\kappa_t|^p] < +\infty$ for any $p \geqslant 1$, such that $|Y_t(x') - Y_t(x)| \leqslant \kappa_t r^\chi |x'-x|^\alpha$, we can bound $A_s(z)$, $\mathbf{P}$ almost surely, by

$$|A_s(z)| \leqslant C'\kappa_t^2 r^{2\chi}|x'-x|^{2\alpha} \int_x^{x'}\int_x^{x'}\int_t^s p_{(s-\rho)/C'}(y-y')\,\mathrm{d}y'\,\mathrm{d}y\,\mathrm{d}\rho \leqslant C'\kappa_t^2 r^{2\chi}(s-t)|x'-x|^{1+2\alpha}.$$

Recalling, for a given $p \geqslant 1$, the standard inequality

$$\begin{aligned}\mathcal{T}_s &:= \int_{\mathbb{R}} \bigl|\partial_x^3 p_{s-t}(z)\bigr| \mathbf{E}\biggl[\biggl(\int_t^s \int_{\mathbb{R}}\biggl(\int_x^{x'} k_{\rho,s}(y-z,u)\bigl[Y_t(y)-Y_t(x)\bigr]\,\mathrm{d}y\biggr)\,\mathrm{d}\zeta(\rho,u)\biggr)^p \bigl|\mathcal{G}_t\biggr]^{1/p}\,\mathrm{d}z \\ &\leqslant C_p \int_{\mathbb{R}} \bigl|\partial_x^3 p_{s-t}(z)\bigr| \mathbf{E}\bigl[\bigl(A_s(z)\bigr)^{p/2}\bigl|\mathcal{G}_t\bigr]^{1/p}\,\mathrm{d}z,\end{aligned}$$

for a universal constant C_p, and using the bound $|\partial_x^3 p_{s-t}(z)| \leqslant C'(s-t)^{-3/2} p_{(s-t)/C'}(z)$, we deduce that (the value of C_p being allowed to increase from line to line) $\mathcal{T}_s \leqslant C_p \kappa_t r^\chi |x-x'|^{1/2+\alpha}(s-t)^{-1}$, so that

$$\mathbf{E}\bigl[|\mathcal{H}_{t,s}(x,x')|^p\bigr]^{1/p} \leqslant C_p \mathbf{E}\bigl[\kappa_t^p\bigr]^{1/p} r^\chi (s-t)^{-1}|x-x'|^{1/2+\alpha}. \tag{77}$$

Since the singularity is non-integrable, we must provide another bound for the left-hand side. By integration by parts in z, we write:

$$\mathcal{H}_{t,s}(x,x') = \int_{\mathbb{R}} \partial_x^2 p_{s-t}(z)\biggl[\int_t^s \int_{\mathbb{R}}\biggl(\int_x^{x'} \partial_z k_{\rho,s}(y-z,u)\bigl[Y_t(y)-Y_t(x)\bigr]\,\mathrm{d}y\biggr)\,\mathrm{d}\zeta(\rho,u)\biggr]\,\mathrm{d}z, \tag{78}$$

which may be split into $\mathcal{H}_{t,s}(x,x') = \mathcal{H}^1_{t,s} + \mathcal{H}^2_{t,s}$, with

$$\begin{aligned}\mathcal{H}^1_{t,s} &= \int_{\mathbb{R}} \partial_x^2 p_{s-t}(z)\biggl[\int_t^s \int_{\mathbb{R}}\biggl(\int_x^{x'} \partial_z k_{\rho,s}(y-z,u)\bigl[Y_t(y)-Y_t(z+u)\bigr]\,\mathrm{d}y\biggr)\,\mathrm{d}\zeta(\rho,u)\biggr]\,\mathrm{d}z, \\ \mathcal{H}^2_{t,s} &= \int_{\mathbb{R}} \partial_x^2 p_{s-t}(z)\biggl[\int_t^s \int_{\mathbb{R}}\biggl(\int_x^{x'} \partial_z k_{\rho,s}(y-z,u)\bigl[Y_t(z+u)-Y_t(x)\bigr]\,\mathrm{d}y\biggr)\,\mathrm{d}\zeta(\rho,u)\biggr]\,\mathrm{d}z.\end{aligned}$$

Repeating the proof of (77), but using in addition the bound for $\partial_z k$ and the fact that $|Y_t(y) - Y_t(z+u)| \leqslant |y-(z+u)|^\alpha$, we obtain

$$\begin{aligned}\mathbf{E}\bigl[|\mathcal{H}^1_{t,s}|^p\bigr]^{1/p} &\leqslant C_p \mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p} r^\chi (s-t)^{-1}|x'-x|\biggl(\int_t^s (\rho-s)^{-1+\alpha}\,\mathrm{d}\rho\biggr)^{1/2} \\ &\leqslant C_p \mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p} r^\chi (s-t)^{-1+\alpha/2}|x'-x|.\end{aligned} \tag{79}$$

(The basic argument is as follows: when computing the bracket of the stochastic integral, the square of gradient of the kernel gives $(r-s)^{-1}$ and the square of $|Y_t(y)-Y_t(z+u)|$ gives $(r-s)^{\alpha}$.)

In order to handle $\mathcal{H}^2_{t,s}$, we notice that

$$\mathcal{H}^2_{t,s}=\int_{\mathbb{R}}\partial_x^2 p_{s-t}(z)\biggl[\int_t^s\int_{\mathbb{R}}\bigl[k_{\rho,s}(x'-z,u)-k_{\rho,s}(x-z,u)\bigr]\bigl[Y_t(z+u)-Y_t(x)\bigr]\,\mathrm{d}\zeta(\rho,u)\biggr]\,\mathrm{d}z.$$

Now, the bracket is less than $C'\kappa_t^2 r^{2\chi}(1+|x'-x|^{\alpha})\int_t^s(s-\rho)^{-1/2}\,\mathrm{d}\rho$, so that

$$\mathbf{E}\bigl[|\mathcal{H}^2_{t,s}|^p\bigr]^{1/p}\leqslant C_p\mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p}r^{\chi}(s-t)^{-3/4}(1+|x'-x|^{\alpha/2}). \tag{80}$$

From (77), (79) and (80) and from a standard interpolation argument, we deduce that, for any $\epsilon\in(0,1/2-\alpha)$, we can find $\varepsilon_1(\epsilon)>0$ and $\varepsilon_2(\epsilon)>0$, both converging to 0 with ϵ, such that

$$\begin{aligned}\mathbf{E}\bigl[|\mathcal{H}_{t,s}|^p\bigr]^{1/p}&\leqslant C'_p\mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p}r^{\chi+\varepsilon_2(\epsilon)}(s-t)^{-1+\varepsilon_1(\epsilon)}|x'-x|^{1/2+\alpha-\epsilon}\\&\leqslant C'_p\mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p}r^{\chi+1/2-\alpha+\varepsilon_2(\epsilon)}(s-t)^{-1+\varepsilon_1(\epsilon)}|x'-x|^{2\alpha},\end{aligned} \tag{81}$$

Recalling that $\alpha>1/3$, we have $1/2-\alpha<\alpha/2$. This says that we can assume without any loss of generality that $\chi>1/2-\alpha$. Then, $r^{\chi+1/2-\alpha+\varepsilon_2(\epsilon)}\leqslant r^{2\chi+\varepsilon_2(\epsilon)}$. By integration from t to T_0, we get

$$\mathbf{E}\bigl[\sup_{t\leqslant T\leqslant T_0}|\mathcal{J}_{t,T}(x,x')|^p\bigr]^{1/p}\leqslant\mathbf{E}\bigl[|\mathcal{K}_t(x,x')|^p\bigr]^{1/p}\leqslant C'_p\mathbf{E}\bigl[|\kappa_t|^p\bigr]^{1/p}r^{2\chi+\varepsilon_2(\epsilon)}|x'-x|^{2\alpha}. \tag{82}$$

Second step. Eq. (82) is the basic step for applying, for a given $0\leqslant t\leqslant T_0$, Kolmogorov's continuity criterion. Anyhow, standard Kolmogorov's criterion doesn't apply since $\mathcal{J}_{t,T}(x,x')$ is not additive in the variables x,x'. To bypass this difficulty, we use Theorem 23 below, which is a refined version of Kolmogorov's criterion. It applies in the current framework because

$$\mathcal{J}_{t,T}(x,y)+\mathcal{J}_{t,T}(y,x')-\mathcal{J}_{t,T}(x,x')=-\bigl(\mathcal{Z}_t^T(x')-\mathcal{Z}_t^T(y)\bigr)\bigl(Y_t(y)-Y_t(x)\bigr).$$

Indeed, assume for the moment that, for any $\alpha'\in(0,\alpha)$, there exists $\varepsilon(\alpha')$, with $\varepsilon(\alpha')\to\varepsilon(\alpha)$ as $\alpha'\to\alpha$, such that $\sup_{0\leqslant T\leqslant T_0}\kappa_{\alpha',\chi+\varepsilon(\alpha')}((\mathcal{Z}_t^T)_{0\leqslant t\leqslant T})$ is in any L^p for any $p\geqslant 1$. Then (87) holds with $\gamma_1=\alpha$, $\gamma_2=\alpha'$ and $\zeta=r^{2\chi+\varepsilon(\alpha')}\kappa'$, where κ' is a random variable with finite moments of any order. The index T here plays the role of the index L in (87): As L is assumed to be in $\mathbb{N}$ in (87), T can be chosen first in a countable dense subset of $[t,T_0]$; as Y_t is assumed to be smooth, cross-integrals can be explicitly defined and there is indeed no difficulty to replace in (82) the supremum over $T\in[t,T_0]$ by a supremum over T in a countable dense subset of $[t,T_0]$.

Applying Theorem 23, we deduce that, for any $\alpha'\in(0,\alpha)$, any $r>0$ and any $0\leqslant t\leqslant T_0$, there exists a random variable $\Gamma_t^{r,\alpha'}$, with finite moments of any order, the bounds being uniform in $t\in[0,T_0]$, such that $\sup_{t\leqslant T\leqslant T_0}|\mathcal{J}_{t,T}(x,x')|\leqslant\Gamma_t^{r,\alpha'}|x-x'|^{2\alpha'}$ for any $x,x'\in[-r,r]$. Pay attention that $\Gamma_t^{r,\alpha'}$ depends on r as Kolmogorov's criterion applies on segments. To get the explicit dependence of the moments of $\Gamma_t^{r,\alpha'}$ upon the variable r, we can apply Kolmogorov's criterion to $\mathcal{J}_{t,T}(rx,rx')$, with $x,x'\in[-1,1]$. We can find $\varepsilon(\alpha')$ (possibly

different from the previous one), with $\varepsilon(\alpha') \to 0$ as $\alpha' \to \alpha$, such that

$$\sup_{r\geqslant 1}\sup_{0\leqslant t\leqslant T_0}\big\{r^{-p[2\chi+\varepsilon(\alpha')]}\mathbf{E}[\sup_{t\leqslant T\leqslant T_0}|\Gamma_t^{r,\alpha'}|^p]\big\}<\infty.$$

Letting $\Gamma_t^{\alpha'}=\sup_{r\in\mathbb{N}^*}\{r^{-[2\chi+2\varepsilon(\alpha')]}\Gamma_t^{r,\alpha'}\}$, we have, for any $p>1/\varepsilon(\alpha')$

$$\sup_{0\leqslant t\leqslant T_0}\mathbf{E}\big[\big(\Gamma_t^{\alpha'}\big)^p\big]\leqslant\sum_{r\in\mathbb{N}^*}\left\{r^{-p[2\chi+2\varepsilon(\alpha')]}\sup_{0\leqslant t\leqslant T_0}\mathbf{E}\big[\big(\Gamma_t^{r,\alpha'}\big)^p\big]\right\}<\infty,$$

which shows that, for any $t\in[0,T_0]$, it holds $\mathbf{P}$-almost surely, for all $x,x'\in[-r,r]$,

$$\sup_{t\leqslant T\leqslant T_0}|\mathcal{J}_{t,T}(x,x')|\leqslant\Gamma_t^{\alpha'}r^{2\chi+2\varepsilon(\alpha')}|x'-x|^{2\alpha'},\tag{83}$$

the random variable $\Gamma_t^{\alpha'}$ having finite moments of any order, the bounds being uniform in $t\in[0,T_0]$.

Third step. The last point of the proof is to 'exchange' the '$\forall t\in[0,T_0]$' and the '$\mathbf{P}$-almost surely' in (83). The argument consists of another application of Kolmogorov's criterion. We thus investigate $\mathcal{J}_{t,T}(x,x')-\mathcal{J}_{t',T}(x,x')$ for $t\leqslant t'\leqslant T$. It can be split into $\int_t^{t'}\mathcal{H}_{t,s}(x,x')\,\mathrm{d}s+\int_{t'}^T[\mathcal{H}_{t,s}(x,x')-\mathcal{H}_{t',s}(x,x')]\,\mathrm{d}s$. By (81), it is quite straightforward to see that

$$\mathbf{E}\left[\left|\int_t^{t'}\mathcal{H}_{t,s}(x,x')\,\mathrm{d}s\right|^p\right]^{1/p}\leqslant C'_p(t'-t)^{\varepsilon_1(\epsilon)}\mathbf{E}\big[|\kappa_t|^p\big]^{1/p}r^{2\chi+\varepsilon_2(\epsilon)}|x'-x|^{2\alpha}.$$

The point is thus to bound $\mathbf{E}[(\int_{t'}^{T_0}|\mathcal{H}_{t,s}(x,x')-\mathcal{H}_{t',s}(x,x')|\,\mathrm{d}s)^p]^{1/p}$. To this end, we notice that $\mathcal{H}_{t,s}(x,x')$ in (76) depends upon t through three quantities: the derivative of the heat kernel, the initial time in the stochastic integral and the increment of Y. Therefore, we must investigate the regularity of each of these three quantities upon t.

We start with the derivative of the heat kernel. Above, we just used the bound $|\partial_x^3p_{s-t}(z)|\leqslant C'(s-t)^{-3/2}p_{(s-t)/C'}(z)$. Noting that we also have $|\partial_s\partial_x^3p_{s-t}(z)|\leqslant C'(s-t)^{-5/2}p_{(s-t)/C'}(z)$, we deduce by interpolation that $\int_{\mathbb{R}}|\partial_x^3p_{s-t}(z)-\partial_x^3p_{s-t'}(z)|\,\mathrm{d}z\leqslant C'(s-t')^{-3/2-\epsilon'}(t'-t)^{\epsilon'}$. This says that, when estimating $\mathbf{E}[|\mathcal{H}_{t,s}(x,x')-\mathcal{H}_{t',s}(x,x')|^p]^{1/p}$, the dependence of the heat kernel upon the time parameter t leads to a slightly modified form of (81), with a new factor $(t'-t)^{\epsilon'}$ and with $\varepsilon_2(\epsilon)$ replaced by $\varepsilon_2(\epsilon)+\epsilon'$.

We now discuss what happens when the initial time varies in the stochastic integral appearing in the definition of $\mathcal{H}_{t,s}(x,x')$. In (77), (79), (80), the interval of integration of the variable ρ is $[t,s]$. This leads to the following singularities: $(s-t)^{-1}$ in (77), $(s-t)^{-1+\alpha/2}$ in (79) and $(s-t)^{-3/4}$ in (80). When $s\geqslant t'$ and ρ lives in $[t,t']$, the length $(s-t)$ can be replaced by $(t'-t)$. This leads, for an arbitrary $\epsilon'\in(0,1)$, to the following new singularities: $(t'-t)^{\epsilon'}(s-t)^{-1-\epsilon'}$ in (77), $(t'-t)^{\epsilon'}(s-t)^{-1-\epsilon'+\alpha/2}$ in (79) and $(t'-t)^{\epsilon'}(s-t)^{-3/4-\epsilon'}$ in (80). Again, this says that, when estimating $\mathbf{E}[|\mathcal{H}_{t,s}(x,x')-\mathcal{H}_{t',s}(x,x')|^p]^{1/p}$, the dependence of the stochastic integral upon the time parameter t leads to a slightly modified form of (81), with a new factor $(t'-t)^{\epsilon'}$ and with $\varepsilon_2(\epsilon)$ replaced by $\varepsilon_2(\epsilon)+\epsilon'$.

We finally discuss what happens when the time parameter varies in the increment of Y. We recall that

$$|Y_{t'}(x)-Y_t(x)|\leqslant\kappa_{t'}r^{\chi'}|t'-t|^\eta,$$

so that (modifying the random variable $\kappa_{t'}$),

$$\big|Y_{t'}(x')-Y_{t'}(x)-\big(Y_t(x')-Y_t(x)\big)\big|\leqslant\kappa_{t'}r^{\chi'}|t'-t|^\eta.$$

Since

$$\big|Y_{t'}(x') - Y_{t'}(x) - \big(Y_t(x') - Y_t(x)\big)\big| \leqslant \big(\kappa_t + \kappa_{t'}\big) r^\chi |x' - x|^\alpha,$$

we get by interpolation (with a new definition of $\kappa_{t'}$)

$$\big|Y_{t'}(x') - Y_{t'}(x) - \big(Y_t(x') - Y_t(x)\big)\big| \leqslant \kappa_{t'} r^{\chi+\epsilon'\chi'} |t'-t|^{\epsilon'\eta} |x'-x|^{(1-\epsilon')\alpha}.$$

This leads to a new version of (77), replacing therein χ by $\chi+\epsilon'\chi'$, α by $(1-\epsilon')\alpha$ and adding the factor $|t'-t|^{\epsilon'\eta}$. For our purpose, there is no need to change (79) and (80) (it is enough to split $Y_{t'}(x')-Y_{t'}(x)-(Y_t(x')-Y_t(x))$ into $Y_{t'}(x')-Y_{t'}(x)$ and $Y_{t'}(x')-Y_{t'}(x)$ and to reproduce the analysis for both terms). This says that, when estimating $\mathbf{E}[|\mathcal{H}_{t,s}(x,x')-\mathcal{H}_{t',s}(x,x')|^p]^{1/p}$, the dependence of Y upon the time parameter t leads to a slightly modified form of (81), with a new factor $(t'-t)^{\epsilon'\eta}$ and with $\varepsilon_2(\epsilon)$ replaced by $\varepsilon_2(\epsilon)+\epsilon'\chi'$ (up to modification of χ') and with α replaced by $\alpha-\epsilon'$.

Collecting the estimates in the three different cases, we end up with

$$\mathbb{E}\Big[\sup_{t'\leqslant T\leqslant T_0} |\mathcal{J}_{t,T}(x,x') - \mathcal{J}_{t',T}(x,x')|^p\Big]^{1/p} \leqslant C'_p r^{2\chi+\varepsilon_2(\epsilon)} |t'-t|^{\varepsilon_3(\epsilon)} |x'-x|^{2\alpha-\epsilon}, \tag{84}$$

with $\varepsilon_2(\epsilon), \varepsilon_3(\epsilon) \to 0$ as $\epsilon \to 0$. The notations in the above inequality are a bit ambiguous: the times t and t' are fixed so that, in the supremum T varies from t' to T_0. In order to avoid such an ambiguity, we let $\mathcal{J}_{t,T}(x,x') = 0$ if $T \leqslant t$. In particular, for $t \leqslant T \leqslant t'$, $|\mathcal{J}_{t,T}(x,x') - \mathcal{J}_{t',T}(x,x')| \leqslant \int_t^{t'} |\mathcal{H}_{t,s}(x,x')| ds$. Reproducing the analysis we just performed, we deduce that the bound right above also holds with $\sup_{t'\leqslant T\leqslant T_0}$ replaced by $\sup_{t\leqslant T\leqslant t'}$ and thus by $\sup_{0\leqslant T\leqslant T_0}$ (as $\mathcal{J}_{t,T}$ and $\mathcal{J}_{t',T}$ vanish when $T \leqslant t$). It says that $\sup_{0\leqslant T\leqslant T_0} |\mathcal{J}_{t,T}(x,x') - \mathcal{J}_{t',T}(x,x')|$ is a random variable, the moments of which are bounded in any L^p in a similar fashion as in the increments in (84). By standard version of Kolmogorov's theorem (applied in the variable t), we deduce that

$$\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T} \sup_{0\leqslant T\leqslant T_0} |\mathcal{J}_{t,T}(x,x')|^p\Big]^{1/p} \leqslant C'_p r^{2\chi+\varepsilon_2(\epsilon)} |x'-x|^{2\alpha-\epsilon}. \tag{85}$$

Now, we can apply the extended version of Kolmogorov's theorem, see Theorem 23 once again, and then complete the proof in the mollified setting.

Fourth step. We now proceed with the last step of the proof. We are to show that the mollified cross-integrals converge almost surely. In that framework, we must recall that the mollification is obtained by a joint mollification of $Y_t(x)$ (in the variable x) and $\mathcal{Y}_{t,s}(x)$ (in the variable x as well) by convolution with the heat kernel. We denote by $(\mathcal{J}^n_{t,T}(x,x'))_{n\geqslant 1}$ the corresponding sequence of cross-integrals, for given values of $t \leqslant T$ in $[0,T_0]$ and $x, x' \in \mathbb{R}$.

The strategy is to discuss the regularity of the sequence $(\mathcal{J}^n_{t,T}(x,x'))_{n\geqslant 1}$ with respect to the mollification parameter n (or more precisely $1/n$, which is the variance of the kernel at rank n). The principle is exactly the same as the one used, in the previous step, for investigating the time regularity. On the same model as (84), we can prove that, for $1 \leqslant m \leqslant n$,

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\leqslant T\leqslant T_0} \big|\big[\big(\mathcal{J}^n_{t',T} - \mathcal{J}^n_{t,T}\big) - \big(\mathcal{J}^m_{t',T} - \mathcal{J}^m_{t,T}\big)\big](x,x')\big|^p\Big]^{1/p} \\ \leqslant C'_p m^{-\varepsilon_3(\epsilon)} r^{2\chi+\varepsilon_2(\epsilon)} |t'-t|^{\varepsilon_3(\epsilon)} |x'-x|^{1/2+\alpha-\epsilon},\end{aligned}$$

with a possibly new value of $\varepsilon_3(\epsilon)$. It is plain to deduce that

$$\mathbb{E}\Big[\sup_{n\geqslant 1} \sup_{0\leqslant T\leqslant T_0} \big|\big(\mathcal{J}^n_{t',T} - \mathcal{J}^n_{t,T}\big)(x,x')\big|^p\Big]^{1/p} \leqslant C'_p r^{2\chi+\varepsilon_2(\epsilon)} |t'-t|^{\varepsilon_3(\epsilon)} |x'-x|^{1/2+\alpha-\epsilon},$$

which is enough to complete the proof.

Final step. To finish with, it remains to check that $(\mathcal{Z}_t^T(x))_{x\in\mathbb{R}}$ satisfies the required regularity assumption, uniformly in $0 \leqslant t \leqslant T \leqslant T_0$. The argument is similar to the one used in the previous step. Indeed, it is quite straightforward to see that

$$\mathcal{Z}_t^T(x) = \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(x-z) \left(\int_t^s \int_{\mathbb{R}} k_{\rho,s}(z,u)\, \mathrm{d}\zeta(\rho,u) \right) \mathrm{d}s\, \mathrm{d}z,$$

so that

$$\begin{aligned}\mathcal{Z}_t^T(x') - \mathcal{Z}_t^T(x) &= \int_t^T \int_{\mathbb{R}} \partial_x^2 p_{s-t}(z) \left[\int_t^s \int_{\mathbb{R}} \big[k_{\rho,s}(x'-z,u) - k_{\rho,s}(x'-z,u) \big]\, \mathrm{d}\zeta(\rho,u) \right] \mathrm{d}s\, \mathrm{d}z \\ &= \int_t^T \int_{\mathbb{R}} \partial_x^3 p_{s-t}(z) \left[\int_t^s \int_{\mathbb{R}} \left(\int_x^{x'} k_{\rho,s}(y-z,u)\, \mathrm{d}y \right) \mathrm{d}\zeta(\rho,u) \right] \mathrm{d}s\, \mathrm{d}z.\end{aligned}$$

The above formula must be compared with (76). It is of the same type but $Y_t(y) - Y_t(x)$ has been replaced by 1. Basically, it says that the above analysis holds but with 2α replaced by α and 2χ replaced by χ. □

5.6. **The environment as the solution of a forward SPDE.** As a concrete application of Propositions 20 and 22, we consider the case when $(Y_t(x))_{0\leqslant t\leqslant T_0, x\in\mathbb{R}}$ is the solution of a forward SPDE of the form

$$\mathrm{d}_t Y_t(x) = L_t Y_t(x)\, \mathrm{d}t + f_t(x)\, \mathrm{d}\zeta(t,x),$$

where $(\zeta(t,x))_{0\leqslant t\leqslant T_0, x\in\mathbb{R}}$ is a Brownian sheet with respect to some filtration $(\mathcal{G}_t)_{0\leqslant t\leqslant T_0}$ on some $(\Xi, \mathcal{G}, \mathbf{P})$, $(L_t = b_t(\cdot)\partial_x + (1/2)\sigma_t^2(\cdot)\partial_x^2)_{0\leqslant t\leqslant T_0}$ is a family of second-order differential operators driven by bounded smooth coefficients with bounded derivatives, the coefficient $(\sigma_t(\cdot))_{0\leqslant t\leqslant T_0}$ being bounded from below by a positive constant, and $(f_t(x))_{0\leqslant t\leqslant T_0, x\in\mathbb{R}}$ is a progressively-measurable random field, uniformly bounded by a deterministic constant (independent of the randomness). The initial condition $(Y_0(x))_{x\in\mathbb{R}}$ is collection of $\mathcal{G}_0$-measurable random variables such that $|Y_0(x) - Y_0(x')| \leqslant \kappa_0 r^\chi |x-x'|^\alpha$, for all $x, x' \in [-r, r]$, with $r \geqslant 1$, where $\alpha \in (1/3, 1/2)$, $\chi \in [0, \alpha/2)$ and κ_0 is a $\mathcal{G}_0$-measurable random variable such that $\mathbb{E}[|\kappa_0|^p] < +\infty$ for all $p \geqslant 1$.

Then, for any $0 \leqslant t \leqslant s \leqslant T_0$, we can express Y_s in terms of the fundamental solution $(q_{t,s}(x,y))_{0\leqslant t<s, x, y\in\mathbb{R}}$ of $(L_t)_{0\leqslant t\leqslant T_0}$, which, together with its derivatives, behave as the Gaussian heat kernel and its derivatives (see [11]). We write

$$Y_s(x) = Y_t(x) + \int_{\mathbb{R}} \big(Y_t(y) - Y_t(x) \big) q_{t,s}(x,y)\, \mathrm{d}y + \int_t^s \int_{\mathbb{R}} f_\rho(y) q_{\rho,s}(x,y)\, \mathrm{d}\zeta(\rho,y). \tag{86}$$

Choosing $t = 0$, we derive a mild representation of $Y_s(x)$ in terms of the initial condition Y_0. Up to a modification of χ, this permits to prove that $(Y_s(x))_{x\in\mathbb{R}}$ satisfies the continuity assumption in Proposition 22. The continuity of the second term in the right-hand side above is easily tackled. The third term can be handled on the same model as $\mathcal{Z}_t^T$ in the proof of Proposition 22.

Then, we can define $(Z_t^T(x))_{x\in\mathbb{R}}$ for any $0 \leqslant t \leqslant T \leqslant T_0$. By the previous results, we can prove that the cross-integral is well-defined. According to the decomposition (86), we can indeed split Z_t^T into three terms: The first one coincides with $\mathcal{Z}_t^T$ in Proposition 18; the second one with $\mathcal{Z}_t^T$ in Proposition 20 and the last one with $\mathcal{Z}_t^T$ in Proposition 22. In each

case, the cross-integral of $\mathcal{Z}_t^T$ with respect to Y_t is well-defined, proving that the cross-integral of Z_t^T with respect to Y_t is also well-defined.

5.7. **Refined Kolmogorov Theorem.** The following theorem is an adaptation of the Kolomorogov continuity theorem, the proof of which is left to the reader:

Theorem 23. *Let $(R_L : [-1,1]^2 \ni (x,y) \mapsto R_L(x,y) \in \mathbb{R})_{L\in\mathbb{N}}$ be a family of random fields such that, for some $p \geqslant 1$, there exist a constant $C > 0$, three exponents $\beta, \gamma_1, \gamma_2 > 0$ and a random variable ζ, satisfying,*

$$\begin{aligned} &\forall x, y \in [-1,1]^2, \quad \mathbb{E}\big[\sup_{L\in\mathbb{N}} |R_L(x,y)|^p\big] \leqslant C|x-y|^{1+\beta}, \\ &\forall L \in \mathbb{N}, \forall x < y < z \in [-1,1], \quad |R_L(x,y) + R_L(y,z) - R_L(x,z)| \leqslant \zeta |x-y|^{\gamma_1}|y-z|^{\gamma_2}. \end{aligned} \tag{87}$$

Then, for $0 < \varsigma < \min(\gamma_1 + \gamma_2, \beta/p)$, there exist constants $C' := C'(\gamma_1, \gamma_2, \beta, p)$ and $C'' := C''(C, \gamma_1, \gamma_2, \beta, p)$ and a non-negative random variable ζ', with $\mathbb{E}[|\zeta'|^p] < C''$, such that,

$$\forall L \in \mathbb{N},\ x < y \in [-1,1], \quad |R_L(x,y)| \leqslant C'\big(\zeta' + \zeta\big)|x-y|^{\varsigma}.$$

6. Connection with the KPZ equation

KPZ equation was introduced by Kardar, Parisi and Zhang in [17] in order to model the growth of a random surface subjected to three phenomena: a diffusion effect, a lateral growth and a random deposit. It has the formal (normalized) shape:

$$\partial_t u(t,x) = \tfrac{1}{2}\partial_x^2 u(t,x) + \tfrac{1}{2}|\partial_x u(t,x)|^2 + \dot{\zeta}(t,x), \tag{88}$$

with 0 as initial condition, where $\dot{\zeta}$ is a time-space white noise (that is the time-space derivative of a Brownian sheet, defined on $(\Xi, \mathcal{G}, \mathbf{P})$ as discussed in Proposition 22). Unfortunately, it is ill-posed. The basic reason is that the gradient is not expected to exist as a true function, thus making the term $|\partial_x u(t,x)|^2$ ill-defined. Formally, such a term must be seen as the square of a distribution.

Two strategies have been developed so far to give a sense to (88). The first one goes back to [4]. It consists in taking benefit of the so-called Hopf-Cole exponential transformation, originally used within the framework of Hamilton-Jacobi-Bellman equations. Basically, u is defined as the exponential of the solution of the stochastic multiplicative heat equation. The second approach is due to Hairer [14] in the case when x is restricted to the torus (in which case ζ is defined accordingly). Therein, the basic point is to solve second-order PDEs driven by a distributional first-order term by means of rough paths theory, which is precisely the strategy we used in Section 3 to solve (12). The two notions coincide but the resulting solution solves a *renormalized* version of (88), which writes (again in a formal sense) as (88) with an additional '$-\infty$' in the right-hand side. The normalization must be understood as follows: When mollifying the noise (say $\dot{\zeta}$ into $\dot{\zeta}^n$), Eq. (88) admits a solution, denoted by u^n, but the sequence $(u^n)_{n\geqslant 1}$ is not expected to converge. To make it converge to the solution of (88), some 'counterterm' must be subtracted to the right-hand side of (88): This counterterm is a constant γ^n depending upon n, which tends to ∞ with n, thus explaining the additional '$-\infty$'.

6.1. **Polymer measure.** For a given $T_0 > 0$ and for any $n \geqslant 1$, we can introduce the (random) polymer measure associated with the noise ζ^n:

$$\frac{d\mathbb{Q}_{\zeta^n}}{d\mathbb{P}} \sim \exp\biggl(\int_0^{T_0} \zeta^n(t, B_t)\,\mathrm{d}t\biggr),$$

where $(B_t)_{0\leqslant t\leqslant T_0}$ is a Brownian motion under $\mathbb{P}$ (independent of ζ), the symbol $\sim$ indicating that the right-hand side is normalized in such a way that $\mathbb{Q}_{\zeta^n}$ is a probability measure. By Girsanov's transformation and Feynman-Kac formula, the dynamics of $(B_t)_{0\leqslant t\leqslant T_0}$ is the solution of the SDE (1) with $Y_t(x) = u^n(T_0 - t, x)$ (the definition of u^n involves γ^n, but γ^n is unnoticeable in the definition of the polymer measure as it is hidden in the normalization constant of the right-hand side). In that framework, our result says that, as n tends to ∞, the law of $(B_t)_{0\leqslant t\leqslant T_0}$ under $\mathbb{Q}_{\zeta^n}$ converges towards the law of the solution $(X_t)_{0\leqslant t\leqslant T_0}$ to (1) with $Y_t(x) = u(T_0 - t, x)$, provided the underlying geometric rough path structure exists, which is checked below in the periodic setting. By [14], the Hölder exponent α in our assumption (see Definition 3) can be chosen as $\alpha = 1/2 - \varepsilon$, ε being as small as desired, and the exponent χ can be chosen as 0 as the solution is periodic. Therefore, the law of $(X_t)_{0\leqslant t\leqslant T_0}$, as defined in Theorem 8, reads as a rigorous interpretation of the (a priori ill-defined) polymer measure $\mathbb{Q}_\zeta$ on the canonical space $\mathcal{C}([0, T_0], \mathbb{R})$, which is a new result.

6.2. **Construction of the rough path.** Generally speaking, the main lines for proving the existence of a geometric cross-integral between $(Y_t(x) = u(T_0 - t, x))_{x\in\mathbb{R}}$ and the corresponding $(Z_t^T(x))_{x\in\mathbb{R}}$, for given $0 \leqslant t \leqslant T \leqslant T_0$, are explained in [14, Subsection 7.1]. Anyhow, the result in [14] doesn't exactly fit our requirements since the version of Z_t^T which is considered therein reads as a stationary version of the one we have been using so far and thus slightly differs from it.

In [14], the rough path structure is shown to exist by means of general results connecting Gaussian processes to rough paths. Actually, our results in Section 5 permit to recover the construction. It is indeed proved in [14, Theorem 1.10] that u expands as the sum of the stationary mean-zero solution u^1 of the stochastic additive heat equation and a Hölder continuous remainder u^2 with a Hölder exponent arbitrarily close to 1. By Proposition 19, the construction of the cross-integral of Z_t^T (with the right definition in (11)) with respect to the remainder $u^2(T_0 - t, \cdot)$ is easily handled. As u can be split into two parts, Z_t^T can be split into two parts as well: One part, denoted by $\mathcal{Z}_t^{T,1}$, involves $u^1(T_0 - t, \cdot)$ and another one, denoted by $\mathcal{Z}_t^{T,2}$, involves $u^2(T_0 - t, \cdot)$. Since the part involving $u^2(T_0 - t, \cdot)$ has the same regularity as $u^2(T_0 - t, \cdot)$, we can apply Proposition 19 again in order to define the cross-integral of $\mathcal{Z}_t^{T,2}$ with respect to $u^1(T_0 - t, \cdot)$.

The only remaining point is thus to construct the cross-integral of $u^1(T_0 - t, \cdot)$ with respect to $\mathcal{Z}_t^{T,1}$. Unfortunately, $u^1(T_0 - \cdot, \cdot)$ is the solution of a backward SPDE whereas results in Subsection 5.6 apply to forward SPDEs, the time reversal affecting the notion of adaptedness. This asks for a rewriting of the argument. The fact that u^1 is required to be of zero-mean (on the torus) doesn't play a role for investigating the existence of the cross-integral as only the space increments of u^1 matter for defining the integral. Therefore, forgetting the zero-mean constraint, we can write, with $Y_t^1(x) = u^1(T_0 - t, x)$,

$$Y_t^1(x) = \int_{\mathbb{S}^1} p_{s-t}(x - y) Y_s^1(y)\,\mathrm{d}y + \int_t^s \int_{\mathbb{S}^1} p_{\rho-t}(x - y)\,\mathrm{d}\zeta(\rho, y),$$

for $0 \leqslant t < s \leqslant T_0$ and $x, y \in \mathbb{S}^1$, where $\mathbb{S}^1$ is the unit circle and $(p_{s-t}(x))_{0\leqslant t<s, x\in\mathbb{R}}$ denotes the heat kernel on $\mathbb{S}^1$. This permits to express the cross integral $\mathcal{I}^1_{t,T}(x,x') := \int_x^{x'} (\mathcal{Z}^{T,1}_t(y) - \mathcal{Z}^{T,1}_t(x))\,\mathrm{d}Y_t(y)$ as (at least in a formal way)

$$\int_x^{x'}\int_t^T\int_{\mathbb{S}^1} \big(\partial_x^2 p_{s-t}(y-z) - \partial_x^2 p_{s-t}(x-z)\big) Y^1_s(z) \bigg(\int_{\mathbb{S}^1} \partial_x p_{s-t}(y-w) Y^1_s(w)\,\mathrm{d}w\bigg)\,\mathrm{d}z\,\mathrm{d}s\,\mathrm{d}y$$
$$+ \int_x^{x'}\int_t^T\int_{\mathbb{S}^1} \big(\partial_x^2 p_{s-t}(y-z) - \partial_x^2 p_{s-t}(x-z)\big) Y^1_s(z) \bigg(\int_t^s\int_{\mathbb{S}^1} \partial_x p_{\rho-t}(y-w)\,\mathrm{d}\zeta(\rho,w)\bigg)\,\mathrm{d}z\,\mathrm{d}s\,\mathrm{d}y.$$

As in the proof of Proposition 22, the second term can be tackled by integration by parts. Transferring the derivative in y from $p_{\rho-t}$ to p_{s-t}, applying stochastic Fubini's theorem (paying attention to the fact that the notion of adaptedness is reversed because of the time-reversal) and then making a change of variable, it reads (up to the boundary terms)

$$-\int_t^T\int_{\mathbb{S}^1} \partial_x^3 p_{s-t}(z) \bigg[\int_t^s\int_{\mathbb{S}^1} \bigg(\int_x^{x'} p_{\rho-t}(y-w)\big(Y^1_s(y-z) - Y^1_s(x)\big)\,\mathrm{d}y\bigg)\,\mathrm{d}\zeta(\rho,w)\bigg]\,\mathrm{d}z\,\mathrm{d}s,$$

which is very close to (76). The proof is then similar to the one of Proposition 22.

The first term in the decomposition of $\mathcal{I}^1_{t,T}(x,x')$ can be split itself into two pieces, according to the symbol 'minus' in the difference $\partial_x^2 p_{s-t}(y-z) - \partial_x^2 p_{s-t}(x-z)$. The first part can be integrated directly. It reads

$$\frac{1}{2}\int_t^T \bigg[\bigg(\int_{\mathbb{S}^1} \partial_x p_{s-t}(x'-z) Y^1_s(z)\,\mathrm{d}z\bigg)^2 - \bigg(\int_{\mathbb{R}} \partial_x p_{s-t}(x-z) Y^1_s(z)\,\mathrm{d}z\bigg)^2\bigg]\,\mathrm{d}s.$$

The second piece can be also computed explicitly. It reads

$$\int_t^T \bigg(\int_{\mathbb{S}^1} \partial_x^2 p_{s-t}(x-z) Y^1_s(z)\,\mathrm{d}z\bigg)\bigg(\int_{\mathbb{S}^1} \big(p_{s-t}(x'-z) - p_{s-t}(x-z)\big) Y^1_s(z)\,\mathrm{d}z\bigg)\,\mathrm{d}s.$$

By standard bounds on p_{s-t}, both terms are bounded by $C|x'-x|^\gamma(\int_t^T (s-t)^{\alpha-1-\gamma/2}\,\mathrm{d}s)$, for any $\gamma < 2\alpha$, which permits to define a 'geometric' cross-integral satisfying (61).

6.3. **Structure of the drift and further prospect.** Decomposition (49) applies and says that the 'drift' part in the dynamics under the polymer measure is almost 3/4-Hölder continuous, which is a new result as well (we refer to [21] for a survey on the connection between KPZ equation and polymers). Actually, as the remainder u^2 in the decomposition of the solution of the KPZ equation has a Hölder exponent close to 1, it can be proved, on the same model as the proof of Remark 15, that only the cross-integral of $u^1(T_0 - t, \cdot)$ with $\mathcal{Z}^{T,1}_t$ matters for computing the cross-integral part in the definition of b in Definition 14.

A challenging question is to investigate the long run behavior of X as T_0 tends to the infinity. Surely, this is connected with the long-run behavior of the solution to the KPZ equation, which is a highly non-trivial question, see again [21] for a survey. This requires a more systematic analysis of the long-run behavior of the solution to (1), which is left as a research program for the future.

References

[1] Amir, G., Corwin, I., and Quastel, J. (2011) Probability distribution of the free energy of the continuum directed random polymer in 1+1 dimensions. *Comm. Pure Appl. Math.*, **64**, 466–537.

[2] Andreoletti, P., and Diel, R. (2011) Limit law of the local time for Brox's diffusion. *J. Theoret. Probab.*, **24**, 634–656.
[3] Bass, R., and Chen, Z.-Q. (2001) Stochastic differential equations for Dirichlet processes. (English summary) *Probab. Theory Related Fields*, **121**, 422–446.
[4] Bertini, L., and Giacomin, G. (1997) Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.*, **183**, 571–607.
[5] Brox, T.(1986) A One-Dimensional Diffusion Process in a Wiener medium. *Ann. Probab.*, **14**, 1206–1218.
[6] Diel, R. (2011) Almost sure asymptotics for the local time of a diffusion in Brownian environment. *Stochastic Process. Appl.*, **121**, 2303-2330.
[7] Ethier, S., and Kurtz, T.G. *Characterization and convergence.* John Wiley & Sons, Inc., New York, 1986.
[8] Flandoli, F. *Random perturbation of PDEs and fluid dynamic models. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010. Lecture Notes in Mathematics, 2015.* Springer, Heidelberg, 2011.
[9] Flandoli, F., Russo, F., and Wolf, J. (2003) Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.*, **40**, 493–542.
[10] Flandoli, F., Russo, F., and Wolf, J. (2004) Some SDEs with distributional drift. II. Lyons-Zheng structure, It's formula and semimartingale characterization. *Random Oper. Stochastic Equations 12*, **2**, 145–184.
[11] Friedman, A. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, Inc., Englewood Cliffs, 1964.
[12] Friz, P., Victoir, N. *Multidimensional Stochastic Processes as Rough Paths. Theory and Applications.* Cambridge University Press, 2010.
[13] Gubinelli, M. (2004) Controlling rough paths. *J. Funct. Anal.*, **216**, 86–140.
[14] Hairer, M. (2013) Solving the KPZ equation. *Ann. of Math.*, **178**, 559–664.
[15] Hu, Y. and Shi, Z. (1998) The Limits of Sinai's Simple Random Walk in Random Environment. *Ann. Probab.*, **26** 1477–1521.
[16] Hu, Y. and Shi Z. (1998) The local time of simple random walk in random environment. *J. Theoret. Probab.*, **11** 765–793.
[17] Kardar, M., Parisi, G., and Zhang, Y.-C. (1986) Dynamical scaling of growing interfaces. *Phys. Rev. Lett.*, **56**, 889–892.
[18] Krylov, N. V., and Röckner, M. (2005) Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, **131**, 154–196.
[19] Lyons, T., Caruana, M., and Lévy, T. *Differential equations driven by rough paths. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, 2004. Lecture Notes in Mathematics, 1908.* Springer, Berlin, 2007.
[20] Lyons, T., and Qian, Z. *System control and rough paths.* Oxford University Press, Oxford, 2002.
[21] Quastel, J. *Introduction to KPZ.* Notes from the Saint-Flour summer school 2012. Available at `http://www.math.toronto.edu/quastel/survey.pdf`
[22] Russo, F., and Trutnau, G. (2007) Some parabolic PDEs whose drift is an irregular random noise in space. *Ann. Probab.*, **35**, 2213–2262.
[23] Stroock, D.W., and Varadhan, S.R.S. *Multidimensional diffusion processes.* Springer-Verlag, Berlin-New York, 1979.
[24] Tanaka, H. (1994) Localization of a diffusion process in a one-dimensional Brownian environment. *Comm. Pure Appl. Math.*, **17**, 755–766.
[25] Veretennikov, A. Yu. (1980) Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb.*, **111**, 434–452.
[26] Young, L.C. (1936) An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, **67**, 251–282.
[27] Zvonkin, A. K. (1974) A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb.*, **93**, 129–149.