EMS Surv. Math. Sci. 13 (2026), 215–354
DOI 10.4171/EMSS/87

A tourist's guide to regularity structures and singular stochastic PDEs

Ismaël Bailleul and Masato Hoshino

Abstract. We give an essentially self-contained treatment of the fundamental analytic and algebraic features of regularity structures and their applications to the study of singular stochastic PDEs.

Contents

1. Introduction

The class of singular stochastic partial differential equations (PDEs) is characterized by the appearance in their formulation of ill-defined products due to the presence in the equation of distributions with low regularity, typically realizations of random distributions. Here are three typical examples.

– The 2- or 3-dimensional parabolic Anderson model equation (PAM)

$$(\partial_t - \Delta_x)u = u\xi, \tag{1.1}$$

Mathematics Subject Classification 2020: 60H15 (primary); 60H17 (secondary).
Keywords: regularity structures, singular stochastic partial differential equations.

with ξ a space white noise. It represents the evolution of a Brownian particle in a 2- or 3-dimensional white noise environment in the torus. (The operator Δ_x stands here for the 2- or 3-dimensional Laplacian.)

– The scalar Φ^4_3 equation from quantum field theory

$$(\partial_t - \Delta_x)u = -u^3 + \zeta, \tag{1.2}$$

with ζ a 3-dimensional spacetime white noise and Δ_x the 3-dimensional Laplacian in the torus or the Euclidean space. Its invariant measure is the scalar Φ^4_3 measure from quantum field theory.

– The generalized (KPZ) equation

$$(\partial_t - \partial_x^2)u = f(u)\zeta + g(u)|\partial_x u|^2, \tag{1.3}$$

with ζ a 1-dimensional spacetime white noise. In a more sophisticated form, it provides amongst others a description of the random motion of a rubber on a Riemannian manifold under a random perturbation of the mean curvature flow motion.

A d-dimensional space white noise has Hölder regularity $-d/2-\kappa$, and a d-dimensional spacetime white noise has Hölder regularity $-d/2-1-\kappa$ under the parabolic scaling, almost surely for every positive κ. Whereas one expects from the heat operator that its inverse regularizes a distribution by 2, this is not sufficient to make sense of any of the products $u\xi$, u^3, $f(u)\zeta$, $|\partial_x u|^2$, $g(u)|\partial_x u|^2$ above, as the product of two Hölder distributions is well defined if and only if the sum of their regularity exponents is positive. Why then bother about such equations? It happens that they appear as scaling limits of a number of microscopic nonlinear random dynamics where the strength of the nonlinearity and the randomness balance each other. Many microscopic random systems exhibit this feature as you will see from reading Corwin and Shen's nice review [30] on singular stochastic PDEs.

A typical statement in the theory of regularity structures [17,20,26,53] about a singular stochastic PDE takes the following informal form, stated here in restricted generality. Consider a subcritical singular stochastic PDE

$$(\partial_t - \Delta_x)u = f(u,\partial u)\zeta + g(u,\partial_x u) =: F(u,\partial u;\zeta) \tag{1.4}$$

driven by a possibly multi-dimensional irregular random noise ζ that is almost surely of spacetime regularity $\alpha-2$, for a deterministic constant $\alpha\in\mathbb{R}$. (The notion of "subcriticality" will be properly defined later in the text.) We talk of the sufficiently regular function F as a "nonlinearity" – even though a particular F could depend linearly or affinely on one or all of its arguments. Denote by $\mathfrak{F}$ the space of nonlinearities that are affine functions of the noise argument. For each $\varepsilon\in(0,1]$, denote also by ζ_ε a regularized version of the noise which converges to ζ as ε goes to 0, obtained, for instance, by convolution $\zeta_\varepsilon := \zeta * \varrho_\varepsilon$ with a deterministic smooth mollifier ϱ_ε. Write

$$u_\varepsilon = \textsc{Sol}(\zeta_\varepsilon; F)$$

for the solution to the well-posed parabolic equation

$$(\partial_t - \Delta_x)u_\varepsilon = F(u_\varepsilon, \partial_x u_\varepsilon; \zeta_\varepsilon)$$

started at time 0 from a given (regular enough) fixed initial condition.

Meta-Theorem 1 (What it means to be a solution). *The following three points hold true.*

- *One can associate to each subcritical singular stochastic PDE a finite-dimensional unbounded Lie group called the renormalization group. Denote by k its generic elements.*
- *This group acts explicitly on the right on the nonlinearity space $\mathfrak{F}$*

$$(k, F) \mapsto F^{(k)} \in \mathfrak{F}. \tag{1.5}$$

- *There exist some deterministic (typically diverging) elements $(k_\varepsilon)_{0<\varepsilon\leq 1}$ of the renormalization group such that, for any element k of the renormalization group, the solutions*

$$\bar{u}_\varepsilon^{(k)} := \mathrm{Sol}\big(\zeta_\varepsilon; (F^{(k_\varepsilon)})^{(k)}\big)$$

to the well-posed stochastic PDE

$$(\partial_t - \Delta_x)\bar{u}_\varepsilon^{(k)} = \big(F^{(k_\varepsilon)}\big)^{(k)}\big(\bar{u}_\varepsilon^{(k)}, \partial_x \bar{u}_\varepsilon^{(k)}; \zeta_\varepsilon\big),$$

with given initial condition, converge in probability in an appropriate function/distribution space to some $\bar{u}^{(k)}$ as ε goes to 0.

A solution to a singular stochastic PDE is not a single function or distribution, but rather the family $(\bar{u}^{(k)})_k$ of functions/distributions indexed by the renormalization group.

The k_ε typically diverge as $\varepsilon > 0$ goes to 0, but there are situations where they remain bounded, or are even constant, like in [19]. We stick to the tradition and talk about any of the above limit functions/distributions as a solution to equation (1.4). We talk of the family of solutions. To have a picture in mind, consider the family of maps

$$S_\varepsilon(x) = (x - 1/\varepsilon)^2 \tag{1.6}$$

on $\mathbb{R}$. It explodes in every fixed interval as ε goes to 0 but remains finite, and converges, in a moving window $S_\varepsilon(x + 1/\varepsilon)$ where it is equal to x^2. It also converges in the other moving window $S_\varepsilon(x + 1/\varepsilon + 1)$ where it is equal to $(x+1)^2$. No given moving window is a priori better than another. In this parallel, the function $\mathrm{Sol}(\zeta_\varepsilon; \cdot)$ plays the role of S_ε, with the infinite-dimensional nonlinearity space $\mathfrak{F}$ in the role of the state space $\mathbb{R}$. The role of the translations $x \mapsto x + 1/\varepsilon$ is played by the group action (1.5) of k_ε on the space of nonlinearities $\mathfrak{F}$. The explicit action of k_ε on the space of nonlinearities gives some formulas of the form

$$F^{(k_\varepsilon)}(u, \partial_x u; \zeta) = f(u, \partial u)\zeta + g(u, \partial_x u) + C_\varepsilon(u, \partial u),$$

for some functions C_ε built from f, g and their derivatives and from k_ε, so

$$\bar{u}_\varepsilon := \textsc{Sol}(\zeta_\varepsilon; F^{(k_\varepsilon)})$$

is the solution to the equation

$$(\partial_t - \Delta_x)\bar{u}_\varepsilon = f(\bar{u}_\varepsilon, \partial_x\bar{u}_\varepsilon)\zeta_\varepsilon + g(\bar{u}_\varepsilon, \partial_x\bar{u}_\varepsilon) + C_\varepsilon(\bar{u}_\varepsilon, \partial_x\bar{u}_\varepsilon), \tag{1.7}$$

with given initial condition. We talk of the function C_ε as a counterterm.

In a robust solution theory for differential equations, a solution to a differential equation ends up being a continuous function of the parameters in the equation. In the case of equation (1.4), the parameters are the functions f, g, the noise ζ, and the initial condition of the equation. While it is unreasonable and wrong to expect that the solutions from Meta-theorem 1 are continuous functions of each realization of the noise, they happen to be continuous functions of a measurable functional of the latter built by probabilistic means. We talk about that functional of the noise as an *enhanced noise*.

Meta-Theorem 2 (Continuity of a solution with respect to the enhanced noise). *For any subcritical singular stochastic PDE* (1.4)*, and for any ζ in a class of random noises including space or spacetime white noises, there is a measurable functional Π of the noise taking values in a metric space such that any individual solution of equation* (1.4) *is a continuous function of* Π.

This is a fundamental point to be compared with the fact that, in the stochastic calculus approach to stochastic (partial) differential equations, the solutions to the equations are only measurable functionals of the noise. In a setting where both approaches can be used and coincide, the regularity structures point of view provides a factorization of the measurable solution map of stochastic calculus under the form of the composition of a measurable function Π of the noise with a continuous function of Π. A number of probabilistic statements about Π are then automatically transferred to the solution of the equation by continuity. The support of the law of the random variable Π determines, for instance, the support of the law of the solution to the equation. A large deviation result for the laws of a family of random Π is also automatically transported by continuity into a large deviation result for the laws of the corresponding family of solutions of the equation.

The functional Π is built as a limit *in probability* of some elementary functionals of the noise. This is in the end the reason why the convergence result in Meta-theorem 1 holds in probability. (The convergence is almost sure along an appropriate sequence.)

How is it that one can prove such statements? The starting point is that solutions of singular stochastic PDEs are not expected to be any kind of Hölder functions or distributions. Rather, under an assumption on the equation captured by the notion of subcriticality, we expect any possible solution

$$u(\cdot) \simeq \sum_\tau u_\tau(x)(\Pi_x\tau)(\cdot) \tag{1.8}$$

to be described locally in terms of a *finite* number of equation-dependent reference functions or distributions $\Pi_x\tau$ that are polynomial functionals of the noise. One of these symbols is denoted by $\mathbf{1}$ and $\Pi_x\mathbf{1}$ is the constant function equal to 1. Unlike u, which may be a distribution in some situations, the u_τ are always some functions. Making the parallel with a classical situation, one could talk of the family $(u_\tau)_\tau$ as a *jet* for u. A fundamental result in regularity structures gives some mild ($(\Pi_x\tau)_{x,\tau}$-dependent) coherence conditions on the $(u_\tau)_\tau$ under which a quantified version of the relation (1.8) determines a unique function/distribution u satisfying it. By trading u for $(u_\tau)_\tau$, the theory of regularity structures then provides a complete description of the local structure of the possible solutions to a given singular stochastic PDE, in terms of their local expansion coefficients with respect to some equation-dependent polynomial functionals $\Pi_x\tau$ of the noise. The theory actually turns the problem upside down by reformulating any singular stochastic PDEs as an equation with unknown some tuple $(u_\tau)_\tau$ of local coefficients satisfying a priori the above-mentioned coherence condition. Since (1.8) gives a local description of u near any state space point x, and $u \simeq u_{\mathbf{1}}(x)$ at first order near x, an equation of the form

$$\mathscr{L}u = F(u;\zeta)$$

will be rewritten near x as

$$\sum_\tau u_\tau(x)\mathscr{L}(\Pi_x\tau) \simeq \sum_k \frac{(\partial_u^k F)(u_{\mathbf{1}}(x))}{k!} \sum_{\tau_1,\dots,\tau_k\neq\mathbf{1}} u_{\tau_1}(x)\cdots u_{\tau_k}(x)(\Pi_x\tau_1)\cdots(\Pi_x\tau_k)\zeta. \tag{1.9}$$

Identifying the terms on both sides provides a triangular system for the $\Pi_x\tau$ and a family of equations for the $u_\tau(x)$. In the example of the parabolic Anderson model equation, one has, for instance, $F(u;\zeta) = u\zeta$, the expansion is indexed by some trees $\tau \in \{\text{♀}, \overset{\circ}{\text{♀}}, \dots\}$, and the system reads

$$\begin{aligned} \mathscr{L}\big(\Pi_x\text{♀}\big) &= \zeta, \\ \mathscr{L}\big(\Pi_x\overset{\circ}{\text{♀}}\big) &= \big(\Pi_x\text{♀}\big)\zeta, \quad \text{etc.} \\ u_\tau(x) &= u_{\mathbf{1}}(x), \quad \text{for all } \tau, \end{aligned}$$

since the nonlinearity is linear here. The base point x in $\Pi_x\tau$ means that, rather than simply defining $\Pi_x\tau$ as $\mathscr{L}^{-1}$ on the right-hand side of its defining equation, we only keep the Taylor remainder at x of this quantity, at some τ-dependent order. Of course, the singular feature of the equation has not disappeared as the products $(\Pi_x\tau_1)\cdots(\Pi_x\tau_k)\zeta$ in (1.9) are still problematic. However, proceeding this way, we have isolated the product problem in the problem of making sense of the reference functions/distributions $\Pi_x\tau$, a task which has little to do with the actual task of solving the equation. This is the very point where the fact that the noise is random plays a crucial role. It allows indeed to build the reference functions/distributions $(\Pi_x\tau)(\omega)$ not as some functions of a realization $\zeta(\omega)$ of the noise but rather as some random variables jointly defined with the noise on a common probability space. This realizes a wonderful decoupling of probability and analysis. *To the former*, we have the task of building the enhanced noise: functions or distribution-valued random variables $\Pi_x\tau$ that involve the noise only. *To the latter*, we have the task of solving uniquely an equation in the side space of local coefficients built from the enhanced noise, regardless of any multiplication problem. The construction of the $\Pi_x\tau$ is done by a limiting procedure called *renormalization*, after similar procedures used in quantum field theory to tackle similar problems.

The fundamentals of the theory of regularity structures were built gradually by M. Hairer and his co-authors in four groundbreaking works [17, 20, 26, 53]. In paper [53], M. Hairer sets the analytic framework of regularity structures and provides an ad hoc study of the renormalization problem for the parabolic Anderson model equation (1.1) and scalar Φ^4_3 equation (1.2). The algebra involved in the renormalization process of a large class of singular stochastic PDEs was unveiled in Bruned, Hairer, and Zambotti's work [20]. The proof that the renormalization algorithm provided in [20] converges was given by Chandra and Hairer in [26]. Last, the fact that the renormalization can be "implemented" at the level of the equation was proved in Bruned, Chandra, Chevyrev, and Hairer's work [17], giving a wonderful analogue of the equivalence of the "subtraction scheme" versus "counterterms" approaches to renormalization problems in quantum field theory. Altogether these four works provide a black box for the local well-posedness theory of subcritical singular stochastic PDEs. This work gives an essentially self-contained short treatment of the fundamental analytic and algebraic features of regularity structures and its applications to the study of singular stochastic PDEs that contains the essential points of the works [17,20,53]. It is intended for readers who already have an idea of the subject and who wish to understand in depth the mechanics at work. We hope nonetheless that even a newcomer to the field may grasp the matter by following the road taken here. Regularity structures and the fundamental tools are developed in generality within a highly abstract setting. No trees are in particular involved in the analysis before we actually construct an example of regularity structure adapted to the study of the generalized (KPZ) equation (1.3) in Section 9. When it comes to applying these tools to singular stochastic PDEs, we trade generality for the concrete example of the generalized (KPZ) equation (1.3), which involves all the difficulties of the most general case. We do not treat Chandra and Hairer's work [26] constructing the measurable functional Π of the noise involved in Meta-theorem 2 using Bruned, Hairer, and Zambotti's renormalization process [20].

We stress here that Hairer's approach to singular stochastic PDEs is somewhat orthogonal to the purely probabilistic approaches of stochastic PDEs pioneered by Pardoux, Walsh, or da Prato and Zabczyk using martingale technics, described in [32] for instance. No knowledge of these approaches is needed to understand what follows.

It is our aim here to give a concise self-contained version of what seems to us to be the most important features of the $433(= 236 + 118 + 79)$ pages of the works [17, 20, 53]. A number of comments about different statements, concepts, and other works are deferred to Appendix D so as to keep focused in the main body of the text. The reader is invited to read this section at any point along their reading. We expect that the reader will see from the present work the simplicity that governs the architecture of the theory. The climb may be hard, but the view after the walk is stunning.

The probabilistic/renormalization side of the analysis of singular stochastic PDEs was not mature yet when this work was first written. This is why the deep work [26] of Chandra and Hairer was left aside in this tourist's guide. The situation has changed after the seminal work [70] of Linares, Otto, Tempelmayr and Tsatsoulis and some subsequent works by Hairer and Steele [61] and Bailleul and Hoshino [11]. We refer the reader to [10] for a review of the subject and further references.

Besides the original articles [17, 20, 26, 53], Hairer's lecture notes [54, 56], the book [43] by Friz and Hairer, Chandra and Weber's article [27], and Berglund's book [15] provide other accessible accounts of part of the material presented here. The work [30] of Corwin and Shen provides a nice non-technical overview of the context in which singular stochastic PDEs arise.

We will introduce the different pieces of the puzzle one after the other to arrive at a clear understanding of the mathematical form of the above meta-theorems. This will be done along the following lines.

1. Concrete regularity structures, models, and modelled distributions. We will first set the scene to talk of the local behaviour of functions/distributions

$$f(\cdot) \sim \sum_{\tau} f_\tau(x)(\Pi_x^{\mathsf{g}}\tau)(\cdot), \tag{1.10}$$

near each spacetime point x, giving a generalization of the notion of jet, in terms of reference functions/distributions $(\Pi_x^{\mathsf{g}}\tau)(\cdot)$ parametrized by τ in a finite set and by the spacetime points x. This will involve the setting of *(concrete) regularity structures*

$$\mathcal{T} = ((T^+, \Delta^+), (T, \Delta))$$

and *models* $\mathsf{M} = (\Pi, g)$, from which the reference functions/distributions $(\Pi_x^{\mathsf{g}}\tau)(\cdot)$ are built. See Sections 2.2 and 2.3 for a detailed description of these objects. The notation $(\Pi_x^{\mathsf{g}}\tau)(\cdot)$ will be defined in Definition 1. In the same way as a family of functions $\{f_k\}_{k\in\mathbb{N}^d}$ on $\mathbb{R}^d$ needs to satisfy a quantitative consistency condition for a function f satisfying

$$f(\cdot) \sim \sum_{k\in\mathbb{N}^d} f_k(x)(\cdot - x)^k, \quad \text{near all } x,$$

to exist, a collection of functions $(f_\tau)_\tau$ needs to satisfy a quantitative consistency condition for a distribution f satisfying (1.10) to exist. This condition will involve the notion of *modelled distribution* and *reconstruction operator* $\mathbf{R}^{\mathsf{M}}$, with a notion of consistency that will depend on the component g of the model M. One can think of a modelled distribution as the consistent jet of a function or a distribution. At that stage, given a regularity structure and a model on it, we will have a convenient way of representing a class of functions/distributions on the state space – not all of them. Given a (system of) singular stochastic PDE(s), a good choice of concrete regularity structure will allow to represent the set of functions/distributions that appear in a naive analysis of the equation via a Picard iteration by some elements of our class. Unlike what happens in the study of controlled ordinary differential equations driven by an ℓ-dimension control, there is no universal concrete regularity structure for the set of all singular stochastic PDEs. One associates to each (system of) subcritical singular stochastic PDE(s) a specific regularity structure.

2. Lifting the equation as an equation in the space of consistent jets. The regularity structure associated with equation (1.4) is built from a noise symbol and some operators $(\mathcal{I}_n)_{n\in\mathbb{N}^{d+1}}$ that play the role in the regularity structure of the operators $\partial^n(\partial_t - \Delta_x)^{-1}$, involved in the Picard fixed-point formulation of the equation. (The letter $\partial = (\partial_t, \partial_x)$ stands here for the time/space derivative operator.) One proceeds then by formulating the equation as a fixed-point problem in the space of consistent jets of functions/distributions $(f_\tau)_\tau$, encoded in the notion of modelled distribution. This will require introducing a tweaked version $\mathcal{K}^{\mathsf{M}}$ of the operator $\mathcal{I}$, as the latter does not produce consistent jets from consistent jets – the notion of consistency depends on g whereas the operator $\mathcal{I}$ does not. The equation on the jet space will happen then to have a unique solution in small time under some mild conditions. This solution will be a continuous function of all the parameters in the equations, the model in particular. Along the way we will turn the initial analytical multiplication problem into the problem of defining some models enjoying some appropriate properties – the so-called *admissible models*. We will see that it is straightforward to construct what is called the canonical lift of a regularized version ζ_ε of the noise ζ as an admissible model M^ε; this can be done for any smooth noise. In those terms, the solution u_ε to a well-posed (system of) singular stochastic PDE(s) driven by some smooth noise ζ_ε can be written as the reconstruction

$$u_\varepsilon = \mathbf{R}^{\mathsf{M}^\varepsilon}(\boldsymbol{u}_\varepsilon) \tag{1.11}$$

of a consistent jet $\boldsymbol{u}_\varepsilon$ obtained as the fixed point

$$\boldsymbol{u}_\varepsilon = \Phi\big(\mathsf{M}^\varepsilon, \boldsymbol{u}_\varepsilon\big) \tag{1.12}$$

of a map Φ that depends continuously on its model argument, so does the reconstruction map $\mathbf{R}^{\mathsf{M}^\varepsilon}$.

3. Renormalized models and renormalized equation. However, the model M^ε does not converge in the appropriate space as the positive regularization parameter ε goes to 0, so a solution to the (system of) singular stochastic PDE(s) under study cannot be defined as the limit of the u_ε as ε goes to 0. The situation is similar to what happens to the function S_ε from (1.6). One has to look at M^ε in a moving window to obtain a finite limit. The *renormalization group* will provide us precisely with this possibility, and will provide us in particular with a family of renormalized canonical models $({}^{k_\varepsilon}\mathsf{M}^\varepsilon)_{0<\varepsilon\leq 1}$. To make the final step from here to the meta-theorems, we will see that this action of the renormalization group on the set of models has a dual action on the space $\mathfrak{F}$ of nonlinearities. The ${}^{k_\varepsilon}\mathsf{M}^\varepsilon$-reconstruction $\bar{u}_\varepsilon$ of the unique ${}^{k_\varepsilon}\mathsf{M}^\varepsilon$-dependent fixed-point equation in the space of jets will happen to solve a "renormalized" version of the singular equation (1.4), with some additional ε-dependent terms diverging most of the time as the regularization parameter ε tends to 0, as in (1.7). The continuity of both the solution of the fixed-point equation (1.12) and the reconstruction map, as functions of the underlying model, will ensure the convergence of $\bar{u}_\varepsilon$ to some limit $\bar{u}$ for some converging renormalized models ${}^{k_\varepsilon}\mathsf{M}^\varepsilon$ with limit $\overline{\mathsf{M}}$, say. The limit function $\bar{u}$ will satisfy a system

$$\bar{u} = \mathbf{R}^{\overline{\mathsf{M}}}(\bar{\boldsymbol{u}}), \quad \bar{\boldsymbol{u}} = \Phi(\overline{\mathsf{M}}, \bar{\boldsymbol{u}}),$$

similar to the system of equations (1.11) and (1.12) satisfied by u_ε. It is in this sense that $\bar{u}$ will deserve to be called a solution of the singular stochastic PDE under study. Think of $\bar{u}$ as a function/distribution defined from its "Taylor" jet $\bar{\boldsymbol{u}}$, with the latter solution of a fixed-point problem.

A word about algebra. It is one of the features of the theory of regularity structures that algebra plays an important role, unlike what one usually encounters in the analytic study of PDEs. This is partly due to the choice of description of the objects involved in the analysis, in terms of "jets-like" quantities. Elementary consistency requirements directly bring algebra into play, under the form of Hopf algebras and actions of the latter on some vector spaces. This is what concrete regularity structures are. The appearance of algebra in the study of singular stochastic PDEs is also due to the fact that the renormalization algorithm used to define the random variables that play the role of a number of ill-defined polynomial functionals of the noise is conveniently encoded in an algebraic structure that we call *renormalization structure*; it differs from a concrete regularity structure. These two points involve Hopf algebras. A last piece of algebra is also needed under the form of some pre-Lie algebras. This is an algebraic structure that behaves as the differentiation operation $(f, g) \mapsto g' f$, derivative of g in the direction of f. Using an algebraic language sheds a gentle light on the meaning of the renormalization process at the level of the equation. Pre-Lie algebras are the ingredient that we use to understand how to build the counterterm C_ε in Meta-theorem 1 and equation (1.7).

The analysis or probability-oriented reader should not be frightened by the perspective of working with some algebraic tools; we will hardly need anything more than a few definitions and elementary facts that are direct consequences of the latter; everything else is proved. We refer the reader to Manchon's lecture notes [73], or the first four chapters of Sweedler's book [78], for some accessible references on Hopf algebras, and to Foissy's work [41] for the basics on pre-Lie algebras; all we need is elementary and recalled below. Appendix B contains in any case all the results from algebra that we use without proving them, with precise pointers to the literature.

Organization of the article. This Tourist's Guide has been organized as follows. Basics on regularity structures are introduced in Section 2, under the form of concrete regularity structures. The reconstruction theorem, which ensures that a consistent jet describes a distribution in the state space, is proved there, in Theorem 4. This allows to formulate in

Section 4 a singular PDE as an equation in a space of modelled distributions over a regularity structure associated with the singular PDE. A fixed-point argument is used in Section 4 to prove a local in time well-posedness result in a space of modelled distributions. Despite their possible differences, the regularity structures built for the study of different subcritical elliptic or parabolic singular PDEs all involve the construction of the counterpart of a (or several) regularizing convolution operator(s) and the proof of its (/their) continuity properties in spaces of modelled distributions. This is done in Section 3. Section 5 sets the scene of renormalization structures. They encode the renormalization algorithm used to build the random variables whose realizations play the role of a finite number of reference functions/distributions. The renormalization algorithm is described in Section 7. The dual action of the renormalization operation on the genuine singular PDE is clarified by the introduction of some pre-Lie structures; this is done in Section 6. Nothing so far requires a deep understanding of how one builds the regularity or the renormalization structure associated with a given singular stochastic PDE. It suffices to assume that they satisfy a small number of simple assumptions to run the analysis. A summary of the assumptions can be found at the beginning of Section 9. Section 9 is dedicated to constructing explicitly such structures in the example of the generalized (KPZ) equation. A summary of our notations is given in Appendix A. Appendix B contains a number of elementary facts from algebra that we use without proof in the text. Precise pointers to the proofs of these facts are given. Appendix C contains the proof of technical results that were not given in the body of the text to keep concentrated on the essential features of the method. A number of comments about the notions, the statements, or the literature are collected in Appendix D. The reader is invited to read them at any time.

Notations. *We use a number of Greek letters with different meanings. As a rule, α, β, γ stand for real numbers, while τ, σ, μ, ν, η, φ, ψ stand for elements of regularity or renormalization structures.*

- *Given two statements $\mathfrak{a}$ and $\mathfrak{a}^+$, we agree to write $\mathfrak{a}^{(+)}$ to mean both the statement $\mathfrak{a}$ and the statement $\mathfrak{a}^+$.*
- *Denote by $e_i = (0,\dots,0,\overset{i}{1},0,\dots,0) \in \mathbb{R}^d$ the i-th basis vector of $\mathbb{R}^d$.*
- *Denote by $\mathbb{N}$ the set of non-negative integers. For each $k = (k_i)_{i=1}^d \in \mathbb{N}^d$ and $x = (x_i)_{i=1}^d \in \mathbb{R}^d$, we use the notations*

$$k! := k_1!\cdots k_d!, \quad x^k := x_1^{k_1}\cdots x_d^{k_d}.$$

For any $k = (k_i)_{i=1}^d$, $\ell = (\ell_i)_{i=1}^d$, we write $\ell \le k$ if $\ell_i \le k_i$ for any i and then define

$$\binom{k}{\ell} := \frac{k!}{\ell!(k-\ell)!} = \binom{k_1}{\ell_1}\cdots\binom{k_d}{\ell_d}.$$

All the notations introduced along the way are gathered in Appendix A, with pointers to the section where they are introduced.

Assumptions. *We emphasize along the way a number of "assumptions" on some regularity structures; they are summarized at the beginning of Section 9. All these properties are satisfied by the regularity structures used to study some singular stochastic PDEs; so, strictly speaking, they are not assumptions. We first present the theory of regularity structures independently of its applications to singular stochastic PDEs and then gradually introduce some more specific features of the particular (tree-indexed) structures that are used in that setting. We find it convenient to state in the form of some "assumption" some special features that a regularity structure can have to emphasize their role in the proofs of some particular results.*

2. Basics on regularity structures

Hairer's theory of regularity structures builds on Gubinelli's approach [48] to T. Lyons' theory of rough paths and rough differential equations [72]. This is a theory of controlled ordinary differential equations

$$dz_t = V(z_t)d\mathbf{X}_t, \quad z_t \in \mathbb{R}^k,$$

driven by irregular controls $\mathbf{X}$. Gubinelli's notion of "path controlled by a rough path $\mathbf{X}$" gives a Taylor-like description of a path around each time s in terms of some "monomials" given by the different components of the increments of the rough path $\mathbf{X}$ between the running time t and the fixed time s. This notion of controlled path turns out to be stable by (regular enough) nonlinear maps and by the operator d^{-1} defining the integral against the reference rough path $\mathbf{X}$. These facts allow to formulate controlled ordinary differential equations driven by a rough path as an integral equation in a space of controlled paths and to prove local well-posedness of the equation under some mild regularity assumptions on the vector fields involved in the dynamics by some fixed-point arguments. (There is no need to know anything about rough path in this work.) Hairer chooses a similar angle to build his theory of singular stochastic PDEs, with an important add-on. In the context of the above controlled differential equation, the theory of regularity structures provides a setting in which one has not only a pointwise description of a potential solution path z but also a "local" description of the $\mathbb{R}^k$-valued distribution $V(z_t)d\mathbf{X}_t$ on some time interval around every fixed time s. This way, all the terms that are involved in the analysis of the controlled differential equation are described in terms of some local expansion.

We start this section by explaining in simple terms why the strategy of local expansion devices automatically brings algebra into play, independently of any problem of dynamical nature. This is the content of Section 2.1. The backbone of these expansion devices is encoded in the notion of concrete regularity structure introduced in Section 2.2. The reference functions/distributions $\Pi_x\tau$ that we informally used in the local expansions (1.8) and (2.1) play the role classically devoted to the Taylor monomials $(\cdot - x)^n$. The $\Pi_x\tau$ are built from two primary objects Π and g that will play in the sequel a crucial role. They define jointly what is called a model. In this setting, the quantification of the approximation (1.8) leads to the notion of modelled distribution. This condition on the collection $(u_\tau(x))_{x,\tau}$ of "coefficients" turns out to guarantee the existence of a function/distribution which is indeed well approximated by the finite sum $\sum_\tau u_\tau(x)\Pi_x\tau$ near any state space point x; a condition for uniqueness of the approximated function/distribution is also known. This fact is the content of the reconstruction theorem. Models, modelled distributions and their reconstruction are presented in Section 2.3. Section 2.4 describes how one can make sense of some nonlinear operations on modelled distributions and gives the properties of a derivative operator acting on the space of modelled distributions.

No PDE-related matters are involved in this section. Its main purpose is to give a general analysis of some arbitrary expansion devices. It is only in Section 3 that we will introduce some particular features of such devices that are involved in the application of this general machinery to the study of singular stochastic PDEs. Together with the introduction of the main objects of regularity structures the main result of this section is the reconstruction theorem, Theorem 4.

2.1. Algebra as the mechanics of local expansion devices

Regularity structures are the backbone of expansion devices for the local description of functions and distributions in (an open set of) a Euclidean space, say $\mathbb{R}^d$. (The isotropic nature of the Euclidean space plays no role here, and as a matter of fact, everything works in a non-isotropic setting. We stick to a Euclidean setting here for simplicity.)

1. The Taylor expansion device. The usual notion of local description of a function near a point $x \in \mathbb{R}^d$ involves the Taylor expansion operation and amounts to comparing a function to a polynomial centered at x

$$f(\cdot) \simeq \sum_n f_n(x)(\cdot - x)^n, \quad \text{near } x. \tag{2.1}$$

The sum over n is finite, the approximation is quantified and we end up describing the class of Hölder functions with real positive regularity exponents. By the binomial expansion, one gets a local description of f near any other point y writing

$$f(\cdot) \simeq \sum_{\ell \le n} f_n(x)\binom{n}{\ell}(\cdot - y)^\ell (y-x)^{n-\ell} \simeq \sum_\ell \Bigg(\sum_{n;\ell\le n} f_n(x)\binom{n}{\ell}(y-x)^{n-\ell}\Bigg)(\cdot - y)^\ell.$$

This expansion brings the important insight that the coefficient $f_\ell(y)$ should be compared with the polynomial centered at x

$$f_\ell(y) \simeq \sum_{n\ge\ell} f_n(x)\binom{n}{\ell}(y-x)^{n-\ell}, \quad \text{near } x, \tag{2.2}$$

where the coefficients $f_n(x)$ are associated with some multi-indices $n \ge \ell$. This relation between the coefficients in the expansion at different points is a *consistency condition*. Conversely, this consistency condition for the coefficients $(f_n)_{n\in\mathbb{N}^d}$ is required to ensure the existence of a function f satisfying (2.1).

2. The reference objects for general expansion devices. A more general local description device involves an $\mathbb{R}^d$-indexed collection of functions or distributions $(\Pi_x \tau)(\cdot)$, with labels τ in a given finite set $\mathcal{B}$. We will consider the functions/distributions locally described as

$$f(\cdot) \simeq \sum_{\tau} f_\tau(x)(\Pi_x \tau)(\cdot), \quad \text{near each } x \in \mathbb{R}^d, \tag{2.3}$$

for some coefficients $f_\tau(x)$. The above expression implicitly assumes that the coefficients $f_\tau(x)$ are function of x. One has $\mathcal{B} = \mathbb{N}^d$ and $(\Pi_x k)(\cdot) = (\cdot - x)^k$, in the usual, Taylor, polynomial setting. So, what is an alternative of the binomial expansion? Like in the former setting, it seems meaningful to impose that $\Pi_x \tau$ is a linear combination of $\Pi_y \sigma$ with several labels σ near another point y. Denoting by T the real vector space spanned by $\mathcal{B}$, we express the situation by the identity

$$(\Pi_x \tau)(\cdot) = \big(\Pi_y(\Gamma_{yx}\tau)\big)(\cdot), \tag{2.4}$$

with a linear map $\Gamma_{yx} : T \to T$. Moreover, since the roles of x and y are exchangeable, the linear maps Γ_{yx} are invertible, with $\Gamma_{yx}^{-1} = \Gamma_{xy}$, and one has a group action of an $\mathbb{R}^d \times \mathbb{R}^d$-indexed group on the local description structure T.

While one uses the same polynomial-type local description (2.2) for the f_n as we do for f in (2.1), in the usual Hölder setting, there is no reason in a more general local description device to use the same reference objects for f and for its local coefficients. This is in particular the case if the $(\Pi_x \tau)(\cdot)$ are meant to describe distributions, among others, while it makes sense to use some functions only as reference objects to describe the functions f_τ. For this reason, we introduce another set $\mathcal{B}^+$ and associate to each $\mu \in \mathcal{B}^+$ with a reference function $\mathrm{g}_{yx}(\mu)$ playing the role of $(y-x)^{n-\ell}$ in (2.2). More precisely, we assume that each f_σ can be compared with a finite linear combination

$$f_\sigma(y) \simeq \sum_{\tau \in \mathcal{B},\, \mu \in \mathcal{B}^+} c_\mu^{\tau,\sigma} f_\tau(x) \mathrm{g}_{yx}(\mu), \quad \text{near } x, \tag{2.5}$$

with some constant $c_\mu^{\tau,\sigma}$ depending on $\tau \in \mathcal{B}, \sigma \in \mathcal{B}^+, \mu \in \mathcal{B}^+$. In the Taylor polynomial setting, one has $\mathcal{B}^+ = \mathcal{B} = \mathbb{N}^d$, $\mathrm{g}_{yx}(k) = (y-x)^k$, and $c_k^{n,\ell} = \mathbf{1}_{k=n-\ell}\binom{n}{\ell}$. In this model situation, all the σ appearing in the above formula have a "size" greater than or equal to the size of τ. Such a hierarchy will be encoded later as a graded structure on T. Unlike the model Taylor polynomial setting, the elements of $\mathcal{B}^+$ are associated with some functions only, while the elements of $\mathcal{B}$ may be associated with some distributions. Note that we take care not to write $(\Pi_x \tau)(y)$ as $\Pi_x \tau$ may be a distribution. It would be consistent to write $\mathrm{g}_x(\mu)(y)$, but we stick to the established and convenient notation $\mathrm{g}_{yx}(\mu)$.

3. Consistency relations. With the above notation, one has

$$f(\cdot) \simeq \sum_{\sigma \in \mathcal{B}} f_\sigma(y)(\Pi_y \sigma)(\cdot) \simeq \sum_{\tau,\sigma \in \mathcal{B},\, \mu \in \mathcal{B}^+} c_\mu^{\tau,\sigma} f_\tau(x) \mathrm{g}_{yx}(\mu)(\Pi_y \sigma)(\cdot).$$

Comparing this expansion to the original expansion (2.3) at the point x, we would have the explicit representation

$$\Gamma_{yx}\tau = \sum_{\sigma\in\mathcal{B},\,\mu\in\mathcal{B}^+} c_\mu^{\tau,\sigma}\, \mathrm{g}_{yx}(\mu)\sigma.$$

To lighten the notation, it will be convenient to introduce a notation under the form of the formal ratio

$$\tau/\sigma = \sum_{\mu\in\mathcal{B}^+} c_\mu^{\tau,\sigma}\mu.$$

It is an element of the real vector space T^+ spanned by $\mathcal{B}^+$. Then, the above consistency formulas read

$$f_\sigma(y) \simeq \sum_{\tau\in\mathcal{B}} f_\tau(x)\mathrm{g}_{yx}(\tau/\sigma), \quad \Gamma_{yx}\tau = \sum_{\sigma\in\mathcal{B}} \mathrm{g}_{yx}(\tau/\sigma)\sigma.$$

We can also derive a transitive relation similar to (2.4) for Γ_{zy} and Γ_{yx}. Indeed, we can develop the coefficient $f_\tau(x)$ in (2.5), and re-indexing the labels, one gets

$$f_\eta(z) \simeq \sum_{\sigma\in\mathcal{B}} f_\sigma(y)\mathrm{g}_{zy}(\sigma/\eta) \simeq \sum_{\sigma,\eta\in\mathcal{B}} f_\tau(x)\mathrm{g}_{zy}(\sigma/\eta)\mathrm{g}_{yx}(\tau/\sigma).$$

In the comparison with (2.5), it would be natural to impose the identity

$$\sum_{\sigma\in\mathcal{B}} \mathrm{g}_{zy}(\sigma/\eta)\mathrm{g}_{yx}(\tau/\sigma) = \mathrm{g}_{zx}(\tau/\eta). \tag{2.6}$$

Together with (2.4), this is a fundamental transition relation analogous to the binomial expansion.

4. Algebra encodes consistency relations. There is a way to describe the skeleton of the relations (2.4) and (2.6) with no mention of any explicit reference functions or distributions. We introduce the splitting maps

$$\Delta : T \to T \otimes T^+, \quad \Delta\tau = \sum_\sigma \sigma \otimes (\tau/\sigma),$$
$$\Delta^+ : T^+ \to T^+ \otimes T^+, \quad \Delta^+(\tau/\eta) = \sum_\sigma (\sigma/\eta) \otimes (\tau/\sigma).$$

(Since there is no unique way to express an element of T^+ in the form τ/σ, the "definition" of $\Delta^+(\tau/\eta)$ may seem problematic. However, this formula actually holds for the "concrete regularity structure" defined below.) Using this notation, the transition map reads

$$\Gamma_{yx} = (\mathrm{Id} \otimes \mathrm{g}_{yx})\Delta,$$

and (2.4) and (2.6) take the form

$$\Pi_x = (\Pi_y \otimes \mathrm{g}_{yx})\Delta, \quad \mathrm{g}_{zx} = (\mathrm{g}_{zy} \otimes \mathrm{g}_{yx})\Delta^+.$$

An important property of these maps is given by the following relations:

$$(\Delta \otimes \mathrm{Id})\Delta = (\mathrm{Id} \otimes \Delta^+)\Delta,$$
$$(\mathrm{Id} \otimes \Delta^+)\Delta^+ = (\Delta^+ \otimes \mathrm{Id})\Delta^+. \tag{2.7}$$

For instance, the transitive relation $\Gamma_{zx} = \Gamma_{zy}\Gamma_{yx}$ is encoded by the first relation as follows:

$$\begin{aligned}\Gamma_{zy}\Gamma_{yx} &= (\mathrm{Id} \otimes \mathrm{g}_{zy} \otimes \mathrm{g}_{yx})(\Delta \otimes \mathrm{Id})\Delta \\ &= (\mathrm{Id} \otimes \mathrm{g}_{zy} \otimes \mathrm{g}_{yx})(\mathrm{Id} \otimes \Delta^+)\Delta = (\mathrm{Id} \otimes \mathrm{g}_{zx})\Delta = \Gamma_{zx}.\end{aligned}$$

Based on the model case of the Taylor expansion device, we ask the family of reference functions $\mathrm{g}_{yx}(\mu)$, $\mu \in \mathcal{B}^+$ to be sufficiently rich to describe locally an *algebra* of functions. The easiest way to ensure that property is to assume that the linear span T^+ of $\mathcal{B}^+$ has an algebra structure and that the maps g_{yx} on T^+ are some characters of this algebra; that is, they are multiplicative maps. We do not impose an algebra structure on T as the elements of $\mathcal{B}$ are meant to be associated with some distributions. In the end, we find that it is natural to introduce a graded vector space T and a graded algebra T^+ with some splitting maps Δ and Δ^+. Using the assumed invertibility of the transition maps Γ_{yx}, an elementary fact from algebra then leads directly to the Hopf algebra structure that appears below in the definition of a concrete regularity structure. (The curious reader can see Proposition 48 in Appendix B. We do not need to understand the details of its simple proof now.)

Note that the dimension d of the state space, or the fact that it is Euclidean, plays no role in this discussion.

We choose to record the essential features of this discussion in the definition of a "concrete" regularity structure given below; this is a special form of the more general notion of regularity structure from Hairer's seminal work [53]. The reader should keep in mind that the entire algebraic setting can be understood at a basic level from the above consistency requirements on a given local description device. We invite the reader to look at Appendix B and read the definitions and basic properties of bialgebras, Hopf algebras, and comodules. It is better to read this appendix in the light of the preceding discussion.

2.2. Regularity structures

We define in this section a particular form of regularity structure that turns out to be sufficient for the study of (systems of) singular stochastic PDE(s). A number of notations are fixed here. A general regularity structure is defined in [53, Section 2] and recalled at the end of this section. The following particular form can be essentially found in [53, Section 8] and [20]. The following definition is to be read in the light of the discussion of Section 2.1 and will be best understood by recalling first from Appendix B the definition of a Hopf algebra and the meaning of connectedness in this setting.

Definition. A *concrete regularity structure* $\mathcal{T} = (T^+, T)$ is the pair of graded vector spaces

$$T^+ = \bigoplus_{\alpha \in A^+} T^+_\alpha, \quad T = \bigoplus_{\beta \in A} T_\beta$$

such that the following holds.

- The vector spaces T^+_α and T_β are finite dimensional.
- The vector space T^+ is a connected graded bialgebra with unit $\mathbf{1}_+$, counit $\mathbf{1}'_+$, coproduct $\Delta^+ : T^+ \to T^+ \otimes T^+$, and grading $A^+ \subset [0, \infty)$.
- The index set A for T is a locally finite subset of $\mathbb{R}$ bounded below. The vector space T is a right comodule over T^+; that is, T is equipped with a splitting map $\Delta : T \to T \otimes T^+$ which satisfies

$$(\Delta \otimes \mathrm{Id})\Delta = (\mathrm{Id} \otimes \Delta^+)\Delta \quad \text{and} \quad (\mathrm{Id} \otimes \mathbf{1}'_+)\Delta = \mathrm{Id}. \tag{2.8}$$

  Moreover, for any $\beta \in A$,

$$\Delta T_\beta \subset \bigoplus_{\alpha \geq 0} T_{\beta-\alpha} \otimes T^+_\alpha. \tag{2.9}$$

We denote by

$$\mathcal{T} := \big((T^+, \Delta^+), (T, \Delta)\big)$$

a concrete regularity structure.

Recall that $\bigoplus$ denotes an algebraic sum, so each element of $T^{(+)}$ is a linear combination of elements of $T^{(+)}_\alpha$ for finitely many α, even if $A^{(+)}$ is an infinite set. The different elements that appear in the definition of a concrete regularity structure will acquire later a concrete meaning. The elements of T and T^+ will index some expansion devices for the study of a given (system of) singular stochastic PDE(s) – remember that each equation will have its own regularity structure. We saw in Section 2.1 the meaning of the splitting maps Δ and Δ^+ and their intertwining/coherence relations (2.7) in terms of some expansion device. Recall from the definition of the graded bialgebra given in Appendix B that $T^+_0 = \langle \mathbf{1}_+ \rangle$ and $T^+_\alpha T^+_\beta \subset T^+_{\alpha+\beta}$ for any $\alpha, \beta \in A^+$. By Proposition 48, the bialgebra T^+ is indeed a Hopf algebra; we denote by S_+ its antipode. Denoting by $\mathcal{M} : T^+ \otimes T^+ \to T^+$ the multiplication operator $\mathcal{M}(a \otimes b) := ab$, and by $\mathbf{1}'_+$ the counit of T^+ – think of it as a dual vector to the vector $\mathbf{1}_+$, the antipode S_+ is characterized by the identity

$$\mathcal{M}(\mathrm{Id} \otimes S_+)\Delta^+\tau = \mathcal{M}(S_+ \otimes \mathrm{Id})\Delta^+\tau = \mathbf{1}'_+(\tau)\mathbf{1}_+.$$

Moreover, the coproduct Δ^+ satisfies $\Delta^+\mathbf{1}_+ = \mathbf{1}_+ \otimes \mathbf{1}_+$, and

$$\Delta^+\tau \in \Big\{\tau \otimes \mathbf{1}_+ + \mathbf{1}_+ \otimes \tau + \sum_{0<\alpha'<\alpha} T^+_{\alpha'} \otimes T^+_{\alpha-\alpha'}\Big\} \tag{2.10}$$

for any $\tau \in T_\alpha^+$ with $\alpha > 0$. Similarly, it is straightforward from (2.8) and (2.9) to check that

$$\Delta\tau \in \left\{\tau \otimes \mathbf{1}_+ + \sum_{\beta'<\beta} T_{\beta'} \otimes T^+_{\beta-\beta'}\right\} \tag{2.11}$$

for any $\tau \in T_\beta$. This identity will later imply for a set $(\Pi_x\tau)_{x,\tau}$ of reference functions/distributions that $\Pi_{x'}\tau \simeq \Pi_x\tau$, up to some terms of smaller "homogeneity" $\beta' < \beta$.

For an arbitrary element τ in T, set

$$\tau = \sum_{\beta\in A} \tau_\beta \in \bigoplus_{\beta\in A} T_\beta.$$

We use a similar notation for elements of T^+. An element τ of $T_\alpha^{(+)}$ is said to be *homogeneous* and is assigned *homogeneity* $|\tau| := \alpha$. The homogeneous spaces T_β and T_α^+ being finite dimensional, all norms on them are equivalent; we use a generic notation $\|\cdot\|_\beta$ or $\|\cdot\|_\alpha$ for norms on these spaces. For simplicity, we write

$$\|\tau\|_\alpha := \|\tau_\alpha\|_\alpha. \tag{2.12}$$

Note that we do not assume any relation between the linear spaces T_α^+ and T_β at that stage. Note also that the homogeneity function $|\cdot|$ takes values in $\mathbb{R}$, and that the parameter β in (2.11) can be non-positive, unlike in (2.10).

Notations.

- *Let $\mathcal{B}_\alpha^+$ and $\mathcal{B}_\beta$ be bases of T_α^+ and T_β, respectively. Set*

$$\mathcal{B}^+ := \bigcup_{\alpha\in A^+} \mathcal{B}_\alpha^+, \quad \mathcal{B} := \bigcup_{\beta\in A} \mathcal{B}_\beta.$$

- *Recall from the end of Section 1 our convention about statements of the form $\mathfrak{s}^{(+)}$. Given $\sigma, \tau \in \mathcal{B}^{(+)}$, we use the notation $\sigma \le^{(+)} \tau$ to mean that $\sigma = \tau$ or $|\sigma| < |\tau|$; we write $\tau/^{(+)}\sigma$ for the element of T^+ defined by the expansion*

$$\Delta^{(+)}\tau = \sum_{\sigma\in\mathcal{B}^{(+)},\, \sigma\le^{(+)}\tau} \sigma \otimes (\tau/^{(+)}\sigma). \tag{2.13}$$

Write $\sigma <^{(+)} \tau$ to mean further that σ is different from τ. The notations $\tau/^{(+)}\sigma$ and $\sigma <^{(+)} \tau$ are only used for τ and σ in $\mathcal{B}^{(+)}$.

It should be noted that, for any regularity structure $\mathcal{T} = (T^+, T)$, the pair

$$\mathcal{T}^+ := \big((T^+, \Delta^+), (T^+, \Delta^+)\big)$$

also define a regularity structure. The polynomial regularity structure defined later takes such a form. Also, an algebraic structure arising from branched rough paths considered in [49] essentially takes the above form. In general, we consider two distinct spaces T^+ and T to encode distributions by the vector space T which is not defined as an algebra.

Interpreting the splitting maps Δ and Δ^+ as chopping elements into pieces, keep in mind that $\tau/^{(+)}\sigma$ can be a sum of elements of T^+, in case σ appears "at different places" as a part of τ. This will be particularly clear in formula (2.14) below for the polynomial regularity structure, where the binomial coefficient $\binom{n}{\ell}$ will account for the number of X^ℓ inside X^n, for $0 \leq \ell \leq n$. Note that, for $\sigma <^{(+)} \tau$ in $\mathcal{B}^{(+)}$, we have

$$\begin{aligned}\Delta^+(\tau/^{(+)}\sigma) &= \sum_{\sigma \leq^{(+)} \eta \leq^{(+)} \tau} (\eta/^{(+)}\sigma) \otimes (\tau/^{(+)}\eta) \\ &= (\tau/^{(+)}\sigma) \otimes \mathbf{1}_+ + \mathbf{1}_+ \otimes (\tau/^{(+)}\sigma) \\ &\quad + \sum_{\sigma <^{(+)} \eta <^{(+)} \tau} (\eta/^{(+)}\sigma) \otimes (\tau/^{(+)}\eta).\end{aligned}$$

These two identities are direct consequences of the *co-associativity properties*

$$(\Delta^{(+)} \otimes \mathrm{Id})\Delta^{(+)} = (\mathrm{Id} \otimes \Delta^+)\Delta^{(+)},$$

of the coproduct $\Delta^{(+)}$, obtained by identifying the corresponding terms on the left- and right-hand sides. In the setting of singular stochastic PDEs where the elements of T are (decorated) trees, τ/σ will be a product of trees, and each of these trees will eventually be involved in the action of re-centering the corresponding analytic objects to a given running point while leaving the trunk tree σ untouched. The definition of a model given in Section 2.3 illustrates exactly this picture.

Here are two examples of regularity structures.

– Let symbols $X_1, \ldots, X_d$ be given. For $n \in \mathbb{N}^d$, set $X^n := X_1^{n_1} \cdots X_d^{n_d}$; this is an element of the free commutative algebra with unit $\mathbf{1}(:= X^0)$ generated by the X_i. We can see that $T_X := \mathrm{span}\{X^n; n \in \mathbb{N}^d\}$ is a bialgebra with the coproduct

$$\Delta_{\mathrm{pol}} X^n := \sum_{\ell \leq n} \binom{n}{\ell} X^\ell \otimes X^{n-\ell}. \tag{2.14}$$

Let $\mathfrak{s} = (\mathfrak{s}_i) \in (\mathbb{N} \setminus \{0\})^d$ be an integer-valued fixed vector, called a *scaling*. This vector accounts for the natural scaling properties in the different directions of $\mathbb{R}^d$ for the problem at hand. If, for instance, $\mathbb{R}^d$ does not stand for the isotropic Euclidean space but rather for a non-isotropic space with topology $\mathbb{R}^d$, as a Lie group, different directions will naturally have different homogeneities, depending on the geometry of the space. We define the scaled degree of $n \in \mathbb{N}^d$ by

$$|n|_{\mathfrak{s}} = \sum_{i=1}^d \mathfrak{s}_i n_i.$$

Then, the definition $T_\alpha = \mathrm{span}\{X^n; |n|_{\mathfrak{s}} = \alpha\}$ gives a grading for the bialgebra T_X. Since $T_0 = \mathrm{span}\{\mathbf{1}\}$, the space T_X is a connected graded bialgebra. Thus, it is indeed a Hopf algebra; the antipode is actually given by $S_+ X^n = (-X)^n$. The *polynomial regularity structure* is given by

$$\mathcal{T}_X := \big((T_X, \Delta_{\text{pol}}), (T_X, \Delta_{\text{pol}})\big).$$

– To have another picture in mind, think of T and T^+ as sets of possibly labelled rooted trees, with T^+ consisting only of trees with positive tree homogeneities – a homogeneity is assigned to each labelled tree. This notion of homogeneity induces the decomposition (2.11) of T into linear spaces spanned by trees with equal homogeneities; a similar decomposition holds for T^+. The coproduct $\Delta^{(+)}\tau$ is typically a sum over subtrees σ of τ with the same root as τ, and τ/σ is the quotient tree obtained from τ by identifying σ with the root; this quotient tree is better seen as a product of trees. See Section 9 for constructions of regularity structures of this sort associated with singular stochastic PDEs. For such regularity structures, the minimum regularity of the elements of T is given by the minimum regularity of the noises in the equation. One can leave aside trees by the time we arrive at Section 9 and work in the abstract setting of this section throughout.

A group of linear transforms on the space T^+ has an important role in [53].

Definition. A *character* g on the Hopf algebra T^+ is a linear map $g : T^+ \to \mathbb{R}$ such that $g(\mathbf{1}_+) = 1$ and $g(\tau_1\tau_2) = g(\tau_1)g(\tau_2)$ for any $\tau_1, \tau_2 \in T^+$. The set G^+ of all characters of the algebra T^+ turns into a group with the convolution product $*$ defined by

$$(g_1 * g_2)\tau := (g_1 \otimes g_2)\Delta^+\tau, \quad \tau \in T^+,$$

where we identify the tensor product of two real numbers with their product. The unit of G^+ is the counit $\mathbf{1}'_+$ of T^+, and the inverse of $g \in G^+$ is given by $g^{-1} := g \circ S_+$, where S_+ is the antipode of T^+. The group G^+ is called *structure group*.

Think of the usual convolution product $(g_1 * g_2)(x) = \int g_1(y)g_2(x-y)dy$, where one first splits x into y and $x - y$, and then applies f and g to each piece, before taking the product and summing over all possible splittings. The group G^+ acts on T from left. One associates to a character g of T^+ the map

$$\hat{g} := (\text{Id} \otimes g)\Delta : T \to T$$

from T to itself. We have indeed

$$\widehat{g_1 * g_2} = \widehat{g_1} \circ \widehat{g_2} \tag{2.15}$$

for any $g_1, g_2 \in G^+$, as a direct consequence of the comodule property (2.8). Indeed,

$$\begin{aligned}
\widehat{g_1 * g_2} &= (\text{Id} \otimes (g_1 * g_2))\Delta = (\text{Id} \otimes g_1 \otimes g_2)(\text{Id} \otimes \Delta^+)\Delta \\
&= (\text{Id} \otimes g_1 \otimes g_2)(\Delta \otimes \text{Id})\Delta = (\widehat{g_1} \otimes g_2)\Delta \\
&= \widehat{g_1} \circ \widehat{g_2}.
\end{aligned}$$

Also, for any $\tau \in T_\beta$,

$$\big(\hat{g}(\tau) - \tau\big) \in \bigoplus_{\beta' < \beta} T_{\beta'},$$

as a consequence of the structural identity (2.11). Similarly, one defines the action of G^+ on T^+ by

$$\hat{g}^+ := (\mathrm{Id} \otimes g)\Delta^+ : T^+ \to T^+$$

for $g \in G^+$. This operator also satisfies the properties similar to $\hat{g}$.

At the end of this section, we recall from [53] the original definition of regularity structures.

Definition. A *regularity structure* $\mathcal{T} = (\mathbf{A}, \mathbf{T}, \mathbf{G})$ consists of the following.

(1) $\mathbf{A}$: a subset of $\mathbb{R}$ such that the set $\{\alpha \in \mathbf{A};\ \alpha < \gamma\}$ is finite for every $\gamma \in \mathbb{R}$.

(2) $\mathbf{T} = \bigoplus_{\alpha \in \mathbf{A}} \mathbf{T}_\alpha$: an algebraic sum of Banach spaces $(\mathbf{T}_\alpha, \|\cdot\|_\alpha)$.

(3) $\mathbf{G}$: a group of continuous linear operators on $\mathbf{T}$ such that, for any $\Gamma \in \mathbf{G}$ and $\alpha \in \mathbf{A}$,

$$(\Gamma - \mathrm{Id})\mathbf{T}_\alpha \subset \bigoplus_{\beta \in \mathbf{A},\, \beta < \alpha} \mathbf{T}_\beta.$$

A concrete regularity structure $((T^+, \Delta^+), (T, \Delta))$ turns into a regularity structure in the above sense, by setting $\mathbf{A} = A$, $\mathbf{T} = T$, and $\mathbf{G} = \{\hat{g}\}_{g \in G^+}$. The map $\hat{g}$ is denoted by Γ_g in Hairer's work [53], where $G = \{\Gamma_g,\ g \in G^+\}$ is defined as a (particular form of) structure group. In this article, we prefer the former Fourier-like notation, which is consistent with the fact that the "hat" map defines a linear representation of G^+ into $L(T)$.

2.3. Models and modelled distributions

The preceding section contains the algebraic backbone of regularity structures. Its analytic flesh is introduced in this section on models and modelled distributions. This analytic setting depends on which (system of) singular stochastic PDE(s) one studies. We will not use the same function spaces to analyze a class of equations involving only the heat operator $(\partial_t - \Delta_x)$ and a system of two equations involving $(\partial_t - \Delta_x)$ for one and an operator with a different scaling for the other, like $(\partial_t + (-\Delta_x)^a)$, or simply $(-\Delta_x)^a$, with $0 < a \neq 1$, for the other. *We choose to concentrate in the present work on parabolic equations involving the heat operator only*. We will thus work throughout with the parabolic space $\mathbb{R} \times \mathbb{R}^d$, with generic point $x = (x_0, x') = (x_i)_{i=0}^d$, equipped with the distance function

$$d(x, y) = d\big((x_0, x'), (y_0, y')\big) = \sqrt{|x_0 - y_0|} + |x' - y'|.$$

The Hölder spaces introduced in the next paragraph of this section will play a prominent role. They are used in the second paragraph to define models over a given regularity structure. Models give us the reference functions/distributions $(\Pi_x\tau)(\cdot)$ and $\mathrm{g}_{yx}(\mu)$ that we will use in our expansion devices to describe potential solutions of given singular stochastic PDEs. Expansion devices associate to each spacetime point a distribution meant to give the local description of a globally defined distribution. There is however no reason that such a globally defined distribution exists if no condition on its local "jets" is imposed. The appropriate consistency condition is encoded in the definition of a modelled distribution. Under this consistency condition, it is a fundamental fact that all these local descriptions can be patched together to define a unique globally defined distribution locally that is close to its local description, everywhere. This is what the reconstruction theorem does for us. We end this section with a paragraph on the special properties of modelled distributions representing functions.

Function spaces. Set

$$\mathfrak{s} := (2, 1, \dots, 1) \in \mathbb{N} \times \mathbb{N}^d,$$

and define, for any multi-index $n = (n_0, n_1, \dots, n_d) \in \mathbb{N} \times \mathbb{N}^d$, the scaled degree of n by

$$|n|_{\mathfrak{s}} := 2n_0 + n_1 + \cdots + n_d.$$

Throughout this article, we define some analytic tools by using the specific kernel approach as introduced in Otto and Weber [76], instead of the local presentation as in [53]. We define a non-positive elliptic operator on $\mathbb{R} \times \mathbb{R}^d$

$$\mathscr{G} := \partial_{x_0}^2 - \Delta_{x'}^2$$

and denote by

$$P_t := e^{t\mathscr{G}}$$

its semigroup, and by $p_t(x, y)$ its kernel with respect to Lebesgue measure. It is a symmetric function of (x, y) that satisfies the scaling property

$$p_t(x, y) = t^{-(d+2)/4} p\Big(\big(t^{-\mathfrak{s}_j/4}(x_j - y_j)\big)_{0 \le j \le d}\Big)$$

for a Schwartz function $p \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$. The estimate

$$\int \big|\partial_x^n p_t(x, y)\big| d^a(x, y) dy \lesssim t^{\frac{a - |n|_{\mathfrak{s}}}{4}} \tag{2.16}$$

holds as a consequence for any multi-index $n \in \mathbb{N} \times \mathbb{N}^d$ and any positive exponent a. For a fixed positive integer $N \in \mathbb{N}$, we define operators $Q_t^{(N)}$ and $P_t^{(N)}$, setting

$$Q_t^{(N)} := (-t\mathscr{G})^N e^{t\mathscr{G}}, \quad P_t^{(N)} := \int_t^\infty Q_s^{(N)} \frac{ds}{s}.$$

This implies that $P_t^{(N)} = \mathsf{P}(-t\mathscr{G}) e^{t\mathscr{G}}$ for a monic polynomial P of degree N. One has in particular $P_t^{(1)} = P_t$, and

$$P_t^{(N)} = \int_t^1 Q_s^{(N)} \frac{ds}{s} + P_1^{(N)}. \tag{2.17}$$

(Those who know a little about Littlewood–Paley decomposition will recognize in $Q_t^{(N)}$ the counterpart of the Littlewood–Paley projectors Δ_i and in the integral with respect to the measure ds/s the counterpart of the uniform measure on the integers; the integral operator associated with $P_1^{(N)}$ plays the role of Δ_{-1}; this is an infinitely smoothing operator.)

Definition. Fix $N \geq 1$ and pick a real number $\alpha < 4N$. We define the *α-Hölder space* $\mathcal{C}^\alpha(\mathbb{R}\times\mathbb{R}^d)$ as the set of tempered distributions on $\mathbb{R}\times\mathbb{R}^d$ with finite $\mathcal{C}^\alpha$-norm defined by

$$\|\Lambda\|_{\mathcal{C}^\alpha} := \big\|P_1^{(N)}(\Lambda)\big\|_{L^\infty(\mathbb{R}\times\mathbb{R}^d)} + \sup_{0<t\leq 1} t^{-\frac{\alpha}{4}}\big\|Q_t^{(N)}(\Lambda)\big\|_{L^\infty(\mathbb{R}\times\mathbb{R}^d)}.$$

We work with the elements of $\mathcal{S}'$ (the space of tempered distributions) satisfying global Hölder estimates. In contrast, Hairer [53] considered the elements of $\mathcal{D}'$ (the dual of compactly supported smooth functions) satisfying local Hölder estimates. Such a technical difference is not important here, but when considering distributions diverging at infinity, the former definition should be appropriately modified by incorporating weight functions. See [11] for instance.

The constraint on α of $\mathcal{C}^\alpha$ comes from the fact that all polynomials of scaled degree no greater than $4N$ are in the kernel of the operator $Q_t^{(N)}$. The above definition of the Hölder spaces depends on N, for the range of regularity exponents considered; write momentarily $\mathcal{C}_N^\alpha(\mathbb{R}\times\mathbb{R}^d)$. We remark that if $\alpha/4 < N < N'$ are given, then one can prove that the N- and N'-dependent norms are equivalent on $\mathcal{C}_N^\alpha(\mathbb{R}\times\mathbb{R}^d)$ – this is a classical fact, worked out, e.g., in [3, Appendix A]. In the sequel, the exponent N is fixed once and for all to a large enough value depending on the problem at hand, so we do not record it in the notations for the Hölder spaces. More generally, one can define Hölder spaces using other elliptic operators than $\mathcal{G}$ with the same "scaling properties" as $\mathcal{G}$; the spaces will be identical and the different norms equivalent. We will use this remark only in the proof of Proposition 11 on the classical Schauder estimates. One can also show that for a positive non-integer regularity exponent a the space $\mathcal{C}^a(\mathbb{R}\times\mathbb{R}^d)$ coincides with the usual space of a-Hölder functions, for the parabolic distance d, with equivalent norms. See, e.g., the proof of [3, Proposition 2.5].

Note that if $\alpha < 0$, then the equivalence

$$\|\Lambda\|_{\mathcal{C}^\alpha} \simeq \sup_{0<t\leq 1} t^{-\frac{\alpha}{4}}\big\|P_t^{(N)}(\Lambda)\big\|_{L^\infty(\mathbb{R}\times\mathbb{R}^d)} \simeq \sup_{0<t\leq 1} t^{-\frac{\alpha}{4}}\|P_t(\Lambda)\|_{L^\infty(\mathbb{R}\times\mathbb{R}^d)} \tag{2.18}$$

holds. The middle term is bounded by the right one, because $P_1^{(1)} = P_1$ and $Q_t^{(1)} = \varphi(t\mathcal{G})P_{t/2}$, with a uniformly bounded operator $\varphi(t\mathcal{G})$. The other direction follows from identity (2.17) relating the operators P and $Q^{(1)}$.

Remark. Otto and Weber [76] were the first to use the semigroup generated by $\mathcal{G}$ in a singular stochastic PDE setting.

Models. Recall from the introduction of Section 2 the intuitive motivation for introducing regularity structures. Whereas the algebra involved in the use of local description devices is captured by the notion of regularity structure, the actual family of functions and distributions involved in these local descriptions is captured by the notion of model over a regularity structure.

Definition 1. A *model over a regularity structure* $\mathcal{T}$ is a pair (g, Π) of maps

$$\mathsf{g} : \mathbb{R}\times\mathbb{R}^d \to G^+, \quad \Pi : T \to \mathcal{S}'(\mathbb{R}\times\mathbb{R}^d)$$

with the following properties.

- Set $\mathsf{g}_{yx} := \mathsf{g}_y * \mathsf{g}_x^{-1}$ for each $x, y \in \mathbb{R}\times\mathbb{R}^d$. For each exponent $\gamma\in\mathbb{R}$, one has

$$\|\mathsf{g}\|_\gamma := \sup_{\tau\in\mathcal{B}^+, |\tau|<\gamma}\ \sup_{x,y\in\mathbb{R}\times\mathbb{R}^d} \frac{|\mathsf{g}_{yx}(\tau)|}{d(y,x)^{|\tau|}} < \infty. \tag{2.19}$$

- The map Π is linear. Set

$$\Pi_x^{\mathsf{g}} := (\Pi\otimes \mathsf{g}_x^{-1})\Delta$$

  for each $x\in\mathbb{R}\times\mathbb{R}^d$. For each exponent $\gamma\in\mathbb{R}$, one has

$$\|\Pi^{\mathsf{g}}\|_\gamma := \sup_{\sigma\in\mathcal{B}, |\sigma|<\gamma}\ \sup_{x\in\mathbb{R}\times\mathbb{R}^d,\, 0<t\le 1} t^{-\frac{|\sigma|}{4}} \left|\left\langle \Pi_x^{\mathsf{g}}\sigma, p_t(x,\cdot)\right\rangle\right| < \infty. \tag{2.20}$$

We also define a pseudo-distance on the space of models over a given regularity structure setting for each $\gamma\in\mathbb{R}$

$$\mathsf{d}_\gamma(\mathsf{M},\mathsf{M}') := \sup_{\substack{\tau\in\mathcal{B}^+, |\tau|<\gamma\\ x,y\in\mathbb{R}\times\mathbb{R}^d}} \frac{|\mathsf{g}_{yx}(\tau) - \mathsf{g}'_{yx}(\tau)|}{d(y,x)^{|\tau|}} + \sup_{\substack{\sigma\in\mathcal{B}, |\sigma|<\gamma\\ x\in\mathbb{R}\times\mathbb{R}^d\\ 0<t\le 1}} t^{-\frac{|\sigma|}{4}} \left|\left\langle \Pi_x^{\mathsf{g}}\sigma - \Pi'^{\mathsf{g}'}_x\sigma, p_t(x,\cdot)\right\rangle\right|. \tag{2.21}$$

By the analytic properties of (g, Π), we have

$$\|\hat{\mathsf{g}}\|_\gamma := \sup_{\sigma,\tau\in\mathcal{B},\, |\sigma|\le|\tau|<\gamma}\ \sup_{x,y\in\mathbb{R}\times\mathbb{R}^d} \frac{|(\widehat{\mathsf{g}_{yx}}(\tau))_\sigma|}{d(y,x)^{|\tau|-|\sigma|}} < \infty, \tag{2.22}$$

where $(\cdot)_\sigma$ denotes the σ-components of elements of T. In Hairer's original work [53], a pair $(\Pi = \{\Pi_x\}_x, \Gamma = \{\Gamma_{yx}\}_{x,y})$ of the family of linear operators

$$\Pi_x : T\to\mathcal{S}'(\mathbb{R}\times\mathbb{R}^d), \quad \Gamma_{yx} : T\to T$$

is called a model if it satisfies analytic conditions (2.20) and (2.22) and algebraic conditions

$$\Pi_x = \Pi_y\Gamma_{yx}, \quad \Gamma_{xx} = \mathrm{Id}, \quad \Gamma_{zx} = \Gamma_{zy}\Gamma_{yx}.$$

In our setting, the choice of operators $\Pi_x := \Pi_x^{\mathrm{g}}$ and $\Gamma_{yx} = \widehat{\mathrm{g}_{yx}}$ provides a model in the original meaning. The algebraic conditions satisfied by Π_x and Γ_{yx} are encoded by the algebraic properties of Δ and Δ^+. Indeed, since $\mathrm{g}_y^{-1} * \mathrm{g}_{yx} = \mathrm{g}_x^{-1}$,

$$\begin{aligned} \Pi_y^{\mathrm{g}} \circ \widehat{\mathrm{g}_{yx}} &= (\Pi \otimes \mathrm{g}_y^{-1} \otimes \mathrm{g}_{yx})(\Delta \otimes \mathrm{Id})\Delta \\ &= (\Pi \otimes \mathrm{g}_y^{-1} \otimes \mathrm{g}_{yx})(\mathrm{Id} \otimes \Delta^+)\Delta \\ &= \big(\Pi \otimes (\mathrm{g}_y^{-1} * \mathrm{g}_{yx})\big)\Delta = \Pi_x^{\mathrm{g}}. \end{aligned} \tag{2.23}$$

So, the above choice is more specific, but such a specific choice can be found in [53, Section 8] and [20, 26]. We also remark that for the analytic estimates in original definition [53] the supremum over x, y are local, and a family compactly supported test functions are considered in the condition (2.20), instead of a single function $p_t(x,\cdot)$. For simplicity, we use the global estimates (2.19) and (2.20) in this article.

Emphasize that g acts on T^+, while Π acts on T, and *note that* g *plays on* T^+ *the same role as that of* Π *on* T: for $\tau \in T^+$ and $\sigma \in T$, one has

$$\mathrm{g}_{yx}(\tau) = \big(\mathrm{g}_y(\cdot) \otimes \mathrm{g}_x^{-1}(\cdot)\big)\Delta^+\tau, \quad (\Pi_x^{\mathrm{g}}\sigma)(y) = \big(\Pi(\cdot)(y) \otimes \mathrm{g}_x^{-1}(\cdot)\big)\Delta\sigma,$$

in a distributional sense for the latter. Therefore, the maps

$$\big(\Pi^{(\mathrm{g})}\tau\big)(x) := \mathrm{g}_x(\tau), \quad x \in \mathbb{R} \times \mathbb{R}^d,\ \tau \in T^+$$

define a model $(\mathrm{g}, \Pi^{(\mathrm{g})})$ on $\mathcal{T}^+ = (T^+, T^+)$.

In the class of problems we consider, it is sufficient in each problem to fix $\gamma \in \mathbb{R}$ to a large enough value; we omit as a consequence this parameter from the notations, unless necessary. We emphasize the dependence of Π_x on g using our notation. We stress that $\Pi\tau$ is only an element of $\mathcal{S}'(\mathbb{R} \times \mathbb{R}^d)$. Think of Π as an interpretation operator for the symbols τ, with τ encoding the structure of the analytic object $\Pi\tau$. One can think of $\Pi_x^{\mathrm{g}}\tau = (\Pi \otimes \mathrm{g}_x^{-1})\Delta\tau$, as $\Pi\tau$ "fully recentered" at x, to give it a concrete meaning. The splitting map Δ identifies the different sets of internal pieces of τ that can be "recentered" to the point x by the action of the map g_x^{-1}, with the full recentering operation on $\Pi\tau$ being the result of all these recentering operations. Condition (2.20) conveys the idea that $\Pi_x^{\mathrm{g}}\tau$ behaves *at point* x like an element of $\mathcal{C}^{|\tau|}(\mathbb{R} \times \mathbb{R}^d)$, as a result of this full recentering operation. We will see in Section 9 concrete examples of recentering operations that can be understood as replacing a function by its Taylor remainder of a certain degree.

The following immediate consequence of the bound (2.20) will be useful in the next section.

Proposition 2. *One has*

$$\sup_{x \in \mathbb{R} \times \mathbb{R}^d,\, 0 < t \le 1} t^{-\frac{|\tau| - |n|_{\mathfrak{s}}}{4}} \big|\big\langle \Pi_x^{\mathrm{g}}\tau, \partial_x^n p_t(x,\cdot)\big\rangle\big| < \infty$$

for any model (Π, g) *on* $\mathcal{T}$, $\tau \in \mathcal{B}$, *and* $n \in \mathbb{N} \times \mathbb{N}^d$.

Proof. By the semigroup property,

$$\partial_x^n p_t(x,y) = \int \partial_x^n p_{\frac{t}{2}}(x,z) p_{\frac{t}{2}}(z,y) dz.$$

We need to apply the distribution $\Pi_x^{\mathrm{g}}\tau$ to the kernel $p_{\frac{t}{2}}(z,\cdot)$. Using the expansion (2.13) of $\Delta\tau$ and the relation (2.23) to write $\Pi_x^{\mathrm{g}}\tau$ in terms of $\Pi_z^{\mathrm{g}}\tau$, one has

$$\left|\left\langle \Pi_x^{\mathrm{g}}\tau, p_{\frac{t}{2}}(z,\cdot)\right\rangle\right| = \left|\sum_{\sigma\le\tau} \mathrm{g}_{zx}(\tau/\sigma)\left\langle \Pi_z^{\mathrm{g}}\sigma, p_{\frac{t}{2}}(z,\cdot)\right\rangle\right| \lesssim \sum_{\sigma\le\tau} d(z,x)^{|\tau|-|\sigma|} t^{\frac{|\sigma|}{4}}$$

from the bound (2.20) in the definition of a model. Using the bound (2.16) on the moments of the heat kernel, we then have

$$\begin{aligned}\left|\left\langle \Pi_x^{\mathrm{g}}\tau, \partial_x^n p_t(x,\cdot)\right\rangle\right| &\lesssim \sum_{\sigma\le\tau} t^{\frac{|\sigma|}{4}} \int \left|\partial_x^n p_{\frac{t}{2}}(x,z)\right| d(z,x)^{|\tau|-|\sigma|} dz \\ &\lesssim \sum_{\sigma\le\tau} t^{\frac{|\sigma|}{4}} t^{\frac{|\tau|-|\sigma|-|n|_{\mathfrak{s}}}{4}} \lesssim t^{\frac{|\tau|-|n|_{\mathfrak{s}}}{4}}.\end{aligned}$$

∎

We close this paragraph by the remark about the situation where all the $\Pi\tau$ are some continuous functions. Then, it follows from the bound on $\langle \Pi_x^{\mathrm{g}}\tau, p_t(x,\cdot)\rangle$, and the fact that $p_t(x,\cdot)$ is converging to a Dirac mass at x, that the function $\Pi_x^{\mathrm{g}}\tau$ satisfies $(\Pi_x^{\mathrm{g}}\tau)(x) = 0$ for all $\tau \in T$ such that $|\tau| > 0$. This will be the case of the smooth (possibly renormalized) models from Section 6.

Modelled distributions and their reconstruction. Think of a T-valued function $\boldsymbol{f}$ on $\mathbb{R}\times\mathbb{R}^d$ as the data needed to associate with each spacetime point $x\in\mathbb{R}\times\mathbb{R}^d$ the local description $\Pi_x^{\mathrm{g}}\boldsymbol{f}(x)$ of a possibly globally defined distribution close to $\Pi_x^{\mathrm{g}}\boldsymbol{f}(x)$ near each x. There is no reason that such a globally defined object exists if one does not impose relations between the different components of $\boldsymbol{f}$. This is what the next definition does. For a real number $\gamma \ge \beta_0$, set

$$T_{<\gamma} := \bigoplus_{\beta<\gamma} T_\beta, \quad T_{<\gamma}^+ := \bigoplus_{\alpha<\gamma} T_\alpha^+.$$

Recall from (2.12) the meaning of the notation $\|\tau\|_\alpha$ for $\alpha\in A$ and $\tau\in T$.

Definition 3. Let $\mathrm{g}:\mathbb{R}\times\mathbb{R}^d \to G^+$ be a function satisfying (2.19). Fix a regularity exponent $\gamma\in\mathbb{R}$. One defines the space $\mathcal{D}^\gamma(T,\mathrm{g})$ of *distributions modelled on the regularity structure* $\mathcal{T}$, *with transition* g, as the space of functions $\boldsymbol{f}:\mathbb{R}\times\mathbb{R}^d \to T_{<\gamma}$ such that

$$\begin{aligned}[\![\boldsymbol{f}]\!]_{\mathcal{D}^\gamma} &:= \max_{\beta<\gamma} \sup_{x\in\mathbb{R}\times\mathbb{R}^d} \left\|\boldsymbol{f}(x)\right\|_\beta < \infty, \\ \|\boldsymbol{f}\|_{\mathcal{D}^\gamma} &:= \max_{\beta<\gamma} \sup_{x,y\in\mathbb{R}\times\mathbb{R}^d} \frac{\|\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\|_\beta}{d(y,x)^{\gamma-\beta}} < \infty.\end{aligned}$$

Set $|||\boldsymbol{f}|||_{\mathcal{D}^\gamma} := [\![\boldsymbol{f}]\!]_{\mathcal{D}^\gamma} + \|\boldsymbol{f}\|_{\mathcal{D}^\gamma}$. We also define the pseudo-distance between two modelled distributions $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$ and $\boldsymbol{f}' \in \mathcal{D}^\gamma(T, \mathrm{g}')$ defined for two distinct models $\mathsf{M} = (g, \Pi)$ and $\mathsf{M}' = (g', \Pi')$, by setting

$$d(\boldsymbol{f}, \boldsymbol{f}') := \sup_{x,y\in\mathbb{R}\times\mathbb{R}^d} \max_{\beta<\gamma} \Bigg\{ \|\boldsymbol{f}(x) - \boldsymbol{f}'(x)\|_\beta + \frac{\|\{\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\} - \{\boldsymbol{f}'(y) - \widehat{\mathrm{g}'_{yx}}\boldsymbol{f}'(x)\}\|_\beta}{d(y,x)^{\gamma-\beta}} \Bigg\}.$$

(As in the definition of models, we choose a global bound to define modelled distributions – the original definition in [53, Section 2] relies on the local bounds. Since we consider the equations on the domain $(x_0, x') \in [0, T] \times \mathbb{T}$, there is no difference between global and local bounds on the spatial variable x'. In Section 4 below, we will consider a weighted norm with respect to the time variable x_0.)

For a basis element $\sigma \in \mathcal{B} \subset T$, and an arbitrary element h in T, denote by h_σ its component on σ in the basis $\mathcal{B}$. For a modelled distribution $\boldsymbol{f}(\cdot) = \sum_{\sigma\in\mathcal{B}} f_\sigma(\cdot)\sigma$ in $\mathcal{D}^\gamma(T, \mathrm{g})$, and $\sigma_0 \in \mathcal{B}$, we have

$$\begin{aligned} \big(\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big)_{\sigma_0} &= f_{\sigma_0}(y) - \sum_{\tau\geq\sigma_0} \mathrm{g}_{yx}(\tau/\sigma_0) f_\tau(x) \\ &= f_{\sigma_0}(y) - f_{\sigma_0}(x) - \sum_{\tau>\sigma_0} \mathrm{g}_{yx}(\tau/\sigma_0) f_\tau(x). \end{aligned} \tag{2.24}$$

Examples.

- The archetype of a modelled distribution is given by the lift

$$\boldsymbol{f}(x) := \sum_{|n|_{\mathfrak{s}}<\gamma} \frac{f^{(n)}(x)}{n!} X^n, \tag{2.25}$$

in the polynomial regularity structure of a γ-Hölder real-valued function f on $\mathbb{R} \times \mathbb{R}^d$ with a positive regularity exponent γ. The identities (2.24) become in that case the Taylor expansions

$$f^{(n)}(y) - f^{(n)}(x) - \sum_{|\ell|_{\mathfrak{s}}<\gamma-|n|_{\mathfrak{s}}} \frac{1}{\ell!} f^{(n+\ell)}(x)(y-x)^\ell = O\big(d(y,x)^{\gamma-|n|_{\mathfrak{s}}}\big) \tag{2.26}$$

satisfied by each $f^{(n)}$. Note here that the function f on $\mathbb{R} \times \mathbb{R}^d$ is γ-Hölder iff there exists a family $(f_n)_{n\in\mathbb{N}\times\mathbb{N}^d}, |n|_{\mathfrak{s}} < \gamma$ of functions on $\mathbb{R} \times \mathbb{R}^d$ satisfying $f_0 = f$ and the condition (2.26) with f_n in the role of $f^{(n)}$. The "if" part holds because one can get $f_n = f^{(n)}$ from (2.26) inductively. The "only if" part is the classical and elementary fact for the isotropic case and can be found in [53, Appendix A] for anisotropic cases. So, the notion of modelled distribution with values in the polynomial regularity structure captures exactly the classical notion of regularity.

- Given a basis element $\tau \in \mathscr{B}$, set

$$\boldsymbol{h}^{\tau}(x) := \widehat{\mathrm{g}_x}\tau - \tau = \sum_{\sigma<\tau} \mathrm{g}_x(\tau/\sigma)\sigma. \tag{2.27}$$

It follows from identity (2.9) in the definition of a concrete regularity structure that $\boldsymbol{h}^{\tau}$ takes values in $T_{<|\tau|}$. Since $\widehat{\mathrm{g}_{yx}} \circ \widehat{\mathrm{g}_x} = \widehat{\mathrm{g}_y}$ by (2.15), it follows that

$$\begin{aligned} \boldsymbol{h}^{\tau}(y) - \widehat{\mathrm{g}_{yx}}\big(\boldsymbol{h}^{\tau}(x)\big) &= (\widehat{\mathrm{g}_y}\tau - \tau) - \widehat{\mathrm{g}_{yx}}(\widehat{\mathrm{g}_x}\tau - \tau) \\ &= (\widehat{\mathrm{g}_y}\tau - \tau) - (\widehat{\mathrm{g}_y}\tau - \widehat{\mathrm{g}_{yx}}\tau) \\ &= \widehat{\mathrm{g}_{yx}}\tau - \tau = \sum_{\sigma<\tau} \mathrm{g}_{yx}(\tau/\sigma)\sigma. \end{aligned}$$

The size estimate $|\mathrm{g}_{yx}(\tau/\sigma)| \lesssim d(y,x)^{|\tau|-|\sigma|}$ required from the g-component of a model then shows that $\boldsymbol{h}^{\tau}$ is a modelled distribution in $\mathcal{D}^{|\tau|}(T_{<|\tau|}, \mathrm{g})$.

- If $\boldsymbol{f}(\cdot) = \sum_{\sigma\in\mathscr{B}} f_\sigma(\cdot)\sigma$ is an element of $\mathcal{D}^{\gamma}(T, \mathrm{g})$, then, for each $\tau \in \mathscr{B}$, the T^+-valued function

$$\boldsymbol{f}/\tau(\cdot) := \sum_{\sigma\geq\tau} f_\sigma(\cdot)\sigma/\tau$$

is an element of $\mathcal{D}^{\gamma-|\tau|}(T^+, \mathrm{g})$, where we denote by $\mathcal{D}^{\gamma}(T^+, \mathrm{g})$ the space of T^+-valued modelled distributions with transition g. Recall that

$$\mathcal{T}^+ = (T^+, T^+)$$

is also a regularity structure.

The next statement says that the consistency condition encoded in the notion of modelled distribution $\boldsymbol{f}$ ensures the existence of a globally defined object close to $\Pi^{\mathrm{g}}_x \boldsymbol{f}(x)$ near each $x \in \mathbb{R}\times\mathbb{R}^d$ and gives condition for uniqueness. Recall that A stands for the index set in the grading of T and set

$$\beta_0 := \min A.$$

Theorem 4 (Reconstruction theorem). *Let $\mathcal{T}$ be a concrete regularity structure and $\mathsf{M} = (\mathrm{g}, \Pi)$ a model over $\mathcal{T}$. Fix a regularity exponent $\gamma \in \mathbb{R}\setminus\{0\}$. There exists a linear continuous operator*

$$\mathbf{R}^{\mathsf{M}} : \mathcal{D}^{\gamma}(T, \mathrm{g}) \to \mathcal{C}^{\beta_0\wedge 0}(\mathbb{R}\times\mathbb{R}^d)$$

satisfying the property

$$\Big|\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi^{\mathrm{g}}_x \boldsymbol{f}(x), p_t(x,\cdot)\big\rangle\Big| \lesssim \|\Pi^{\mathrm{g}}\| \, \|\boldsymbol{f}\|_{\mathcal{D}^{\gamma}} t^{\frac{\gamma}{4}}, \tag{2.28}$$

uniformly in $\boldsymbol{f} \in \mathcal{D}^{\gamma}(T, \mathrm{g})$, $x \in \mathbb{R}\times\mathbb{R}^d$, and $0 < t \leq 1$. Such an operator is unique if the exponent γ is positive.

A distribution $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}$ satisfying identity (2.28) is called a *reconstruction of the modelled distribution* $\boldsymbol{f}$. When $\gamma = 0$, the existence of a reconstruction is not ensured by (2.28) in general. See [23, Example 5.5]. We will see as a particular case of Corollary 6 that the lift (2.25) of a γ-Hölder function f in the polynomial regularity structure has indeed f as a reconstruction.

Notice that from the definition of Π_x^{g}, we have the relation

$$\begin{aligned}(\Pi_x^{\mathsf{g}} \otimes \mathsf{g}_x)\Delta &= (\Pi \otimes \mathsf{g}_x^{-1} \otimes \mathsf{g}_x)(\Delta \otimes \mathrm{Id})\Delta \\ &= (\Pi \otimes \mathsf{g}_x^{-1} \otimes \mathsf{g}_x)(\mathrm{Id} \otimes \Delta^+)\Delta \\ &= (\Pi \otimes \mathbf{1}'_+)\Delta = \Pi,\end{aligned}$$

and thus,

$$\Pi\tau = \Pi_x^{\mathsf{g}}(\widehat{\mathsf{g}_x}\tau) = \Pi_x^{\mathsf{g}}\tau + \Pi_x^{\mathsf{g}}\boldsymbol{h}^\tau.$$

Therefore, the constraint $|\langle \Pi_x^{\mathsf{g}}\tau, p_t(x,\cdot)\rangle| \lesssim t^{|\tau|/4}$, which needs to be satisfied by a model, is equivalent to the estimate

$$\Big|\big\langle \Pi\tau - \Pi_x^{\mathsf{g}}\boldsymbol{h}^\tau(x), p_t(x,\cdot)\big\rangle\Big| = \Big|\Big\langle \Pi\tau - \sum_{\sigma<\tau} \mathsf{g}_x(\tau/\sigma)\Pi_x^{\mathsf{g}}\sigma, p_t(x,\cdot)\Big\rangle\Big| \lesssim t^{|\tau|/4}, \qquad (2.29)$$

which says that $\Pi\tau$ is a/the reconstruction of the modelled distribution $\boldsymbol{h}^\tau$ from (2.27), depending on whether $|\tau| \le 0$ or $|\tau| > 0$. Since, for $|\tau| < 0$, the difference $(*)$ of two reconstructions of $\boldsymbol{h}^\tau$ satisfies

$$\big|\big\langle (*), p_t(x,\cdot)\big\rangle\big| \lesssim t^{|\tau|/4}$$

for all $x \in \mathbb{R}\times\mathbb{R}^d$, this difference is a $\mathcal{C}^{|\tau|}(\mathbb{R}\times\mathbb{R}^d)$ distribution from identity (2.18). So, the estimate (2.29) shows in particular that we could require from scratch that the Π map of a model of $\mathcal{T}$ takes values in $\mathcal{C}^{\beta_0\wedge 0}(\mathbb{R}\times\mathbb{R}^d)$ rather than $\mathcal{S}'(\mathbb{R}\times\mathbb{R}^d)$. The case $|\tau| = 0$ does not cause any problem as we assume that the only element of T of null homogeneity is $\mathbf{1}$.

We will only work with $\mathcal{D}^\gamma(T,\mathsf{g})$-spaces with positive regularity exponents γ in our study of singular stochastic PDEs. We only give a proof of the reconstruction theorem in that setting, following Otto and Weber's nice approach [76]. An extension to the inhomogeneous integral kernels can be found in [64]. See Friz and Hairer's lecture notes [43] for another treatment along these lines. See Hairer's original work [53] or the references given in Appendix D for a proof of Theorem 4 in the case $\gamma \le 0$.

Proof. Existence. We construct explicitly a reconstruction operator. Note first that since

$$\begin{aligned}\big(\Pi_y^{\mathsf{g}}\boldsymbol{f}(y) - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x)\big)(\cdot) &= \big(\Pi_y^{\mathsf{g}}\big(\boldsymbol{f}(y) - \widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big)\big)(\cdot) \\ &= \sum_{\tau\in\mathcal{B}}\big(\boldsymbol{f}(y) - \widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big)_\tau\big(\Pi_y^{\mathsf{g}}\tau\big)(\cdot),\end{aligned}$$

one has

$$\left|\left\langle \Pi^{\mathsf g}_y \boldsymbol f(y) - \Pi^{\mathsf g}_x \boldsymbol f(x), p_t(y,\cdot)\right\rangle\right| \lesssim \sum_{\tau\in\mathcal B, |\tau|<\gamma} d(y,x)^{\gamma-|\tau|} t^{\frac{|\tau|}{4}},$$

from the bounds on models and modelled distributions. For $0<s\le t\le 1$ and $x\in\mathbb R\times\mathbb R^d$, set

$$\mathbf I^t_s(x) := \int p_{t-s}(x,y)\left\langle \Pi^{\mathsf g}_y \boldsymbol f(y), p_s(y,\cdot)\right\rangle dy.$$

We will obtain the distribution $\mathsf R^{\mathsf M}\boldsymbol f$ from $\mathbf I^t_s$ under the form $\lim_{t\downarrow 0}\lim_{s\downarrow 0}\mathbf I^t_s$, with the limits taken in that order, with s sent to 0 first and then t sent to 0. First, from the bounds on modelled distributions, we have

$$\mathbf I^t_t(x) = \left\langle \Pi^{\mathsf g}_x \boldsymbol f(x), p_t(x,\cdot)\right\rangle \quad\text{and}\quad \left|\mathbf I^t_t(x)\right| \le \sum_{\tau\in\mathcal B}\left|f_\tau(x)\right|\left|\left\langle \Pi^{\mathsf g}_x \tau, p_t(x,\cdot)\right\rangle\right| \lesssim t^{\frac{\beta_0}{4}},$$

and moreover, for $0<s'<s<t\le 1$, we have from the semigroup property of the kernel p the x-uniform estimate

$$\begin{aligned}\left|\mathbf I^t_{s'}(x) - \mathbf I^t_s(x)\right| &= \left|\int p_{t-s}(x,z)p_{s-s'}(z,y)\left\langle \Pi^{\mathsf g}_y \boldsymbol f(y) - \Pi^{\mathsf g}_z \boldsymbol f(z), p_{s'}(y,\cdot)\right\rangle dz\,dy\right| \\ &\le \sum_{\tau\in\mathcal B, |\tau|<\gamma}\int p_{t-s}(x,z)p_{s-s'}(z,y) d(y,z)^{\gamma-|\tau|}(s')^{\frac{|\tau|}{4}} dz\,dy \\ &\lesssim \sum_{\tau\in\mathcal B, |\tau|<\gamma} (s-s')^{\frac{\gamma-|\tau|}{4}}(s')^{\frac{|\tau|}{4}}.\end{aligned}$$

For $s'\in[s/2,s)$, this implies that

$$\left|\mathbf I^t_{s'}(x) - \mathbf I^t_s(x)\right| \lesssim s^{\frac{\gamma}{4}}. \tag{2.30}$$

For $s'\in(0,s/2)$, by taking $n\in\mathbb N$ such that $s'\in[s/2^{n+1}, s/2^n)$, we have

$$\begin{aligned}\left|\mathbf I^t_{s'}(x) - \mathbf I^t_s(x)\right| &\le \sum_{m=0}^{n-1}\left|\mathbf I^t_{s/2^m}(x) - \mathbf I^t_{s/2^{m+1}}(x)\right| + \left|\mathbf I^t_{s/2^n}(x) - \mathbf I^t_{s'}(x)\right| \\ &\lesssim \sum_{m=0}^{n-1}(s/2^m)^{\frac{\gamma}{4}} + (s/2^n)^{\frac{\gamma}{4}} \lesssim s^{\frac{\gamma}{4}}.\end{aligned}$$

Thus, the bound (2.30) holds uniformly over $0<s'<s$. Hence, the (locally in t) uniform limit

$$\mathbf I^t_0(x) := \lim_{s\to 0}\mathbf I^t_s(x)$$

exists, since γ is positive. As the identity $P_{t'}\mathbf I^t_0 = \mathbf I^{t+t'}_0$ follows from the semigroup property, we see from (2.18) that $\{\mathbf I^t_0\}_{0<t\le 1}$ is bounded in the space $\mathcal C^{\beta_0}(\mathbb R\times\mathbb R^d)$. (Note that all of the above estimates on $\mathbf I^t_s$ hold over $0<s\le t\le 2$, since the bounds on $\Pi^{\mathsf g}_x\tau$ can be

extended to $0 < t \leq 2$ by a similar argument to Proposition 2.) Therefore, by noting the continuity of $t \mapsto P_t\Lambda$,

$$\|(P_t - \mathrm{Id})\Lambda\|_{\mathcal{C}^{\beta_0-\varepsilon}} \lesssim t^{\frac{\varepsilon}{4}}\|\Lambda\|_{\mathcal{C}^{\beta_0}}$$

for any $\varepsilon > 0$ and $\Lambda \in \mathcal{C}^{\beta_0}$ (see, e.g., [64, Lemma 2.15]) for $0 < s < t \leq 1$, we have

$$\|\mathbf{I}_0^t - \mathbf{I}_0^s\|_{\mathcal{C}^{\beta_0-\varepsilon}} = \|(P_{t-s} - \mathrm{Id})\mathbf{I}_0^s\|_{\mathcal{C}^{\beta_0-\varepsilon}} \lesssim (t-s)^{\frac{\varepsilon}{4}}\|\mathbf{I}_0^s\|_{\mathcal{C}^{\beta_0}} \lesssim (t-s)^{\frac{\varepsilon}{4}}.$$

Hence, $\{\mathbf{I}_0^t\}_{0<t\leq 1}$ converges in $\mathcal{C}^{\beta_0-\varepsilon}(\mathbb{R}\times\mathbb{R}^d)$ as t goes to 0 for any $\varepsilon > 0$. Denote its limit by $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}$. Since

$$\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, p_t(x,\cdot)\big\rangle = \lim_{s\to 0}\big\langle \mathbf{I}_0^s, p_t(x,\cdot)\big\rangle = \lim_{s\to 0}\mathbf{I}_0^{t+s}(x) = \mathbf{I}_0^t(x),$$

we have actually $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} \in \mathcal{C}^{\beta_0}(\mathbb{R}\times\mathbb{R}^d)$ from the x-uniform bound $|\mathbf{I}_0^t(x)| \lesssim t^{\beta_0/4}$. Letting $s = t$ and sending s' to 0 in (2.30), we can check that $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}$ satisfies the bound (2.28).

Uniqueness. To prove the uniqueness of the reconstruction operator on $\mathcal{D}^\gamma(T,\mathrm{g})$ when the regularity exponent γ is positive, we start from the identity

$$\big|\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\big\rangle\big| \lesssim t^{\frac{\gamma}{4}},$$

satisfied uniformly in $x \in \mathbb{R}\times\mathbb{R}^d$ by any other reconstruction operator $(\mathbf{R}^{\mathsf{M}})'$. As for any Schwartz function $\varphi \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^d)$ the convolutions $\int \varphi(x)p_t(x,z)dx$ converge to φ in the smooth topology, one has from the symmetry of the kernels p_t and the fact that γ is positive

$$\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \big(\mathbf{R}^{\mathsf{M}}\big)'\boldsymbol{f}, \varphi\big\rangle = \lim_{t\to 0}\int \big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\big\rangle\varphi(x)dx = \lim_{t\to 0} O\big(t^{\frac{\gamma}{4}}\big) = 0. \qquad \blacksquare$$

One can use Proposition 2 to improve estimate (2.28) under the form

$$\big|\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x), \partial_x^n p_t(x,\cdot)\big\rangle\big| \lesssim t^{\frac{\gamma-|n|_{\mathfrak{s}}}{4}}, \tag{2.31}$$

uniformly in $x \in \mathbb{R}\times\mathbb{R}^d$, for each $n \in \mathbb{N}\times\mathbb{N}^d$.

We now state two standard properties of the reconstruction operator. The following fact implies that $\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, \varphi\rangle$ depends only on the restriction of $\boldsymbol{f}$ to the support of φ. This fact is used to define the reconstructions of modelled distributions which are given in $(0,t)\times\mathbb{R}^d$ with $t \in (0,\infty]$, not in $\mathbb{R}\times\mathbb{R}^d$. See Theorem 20 and Section 4.3.

Corollary 5. *Pick γ positive. If $\boldsymbol{f} \in \mathcal{D}^\gamma(T,\mathrm{g})$ is null on an open set $U \subset \mathbb{R}\times\mathbb{R}^d$, then $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} = 0$ on U.*

Proof. Since the mapping $x \mapsto \Pi_x^{\mathrm{g}}\boldsymbol{f}(x)$ is null on U, it follows from estimate (2.28) that

$$\big|\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, p_t(x,\cdot)\big\rangle\big| \lesssim t^{\gamma/4}$$

for all $x \in U$. For a smooth function φ with compact support in U, one can use the convergence of $\int \varphi(x) p_t(x,z)dx$ to φ as $t > 0$ goes to 0 in an appropriate C^k space to get

$$\big\langle \mathbf{R}^{\mathsf{M}} \boldsymbol{f}, \varphi \big\rangle = \lim_{t\to 0} \int \varphi(x) \big\langle \mathbf{R}^{\mathsf{M}} \boldsymbol{f}, p_t(x,\cdot) \big\rangle dx = 0. \qquad \blacksquare$$

The following fact is an immediate consequence of uniqueness in the reconstruction theorem; it is used in Section 4 and implies in particular that the lift (2.25) of a γ-Hölder function f in the polynomial regularity structure has indeed f as a reconstruction.

Corollary 6. *Pick γ positive and $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$. If the model (g, Π) takes values in the space of smooth functions on $\mathbb{R} \times \mathbb{R}^d$, then the mapping*

$$x \mapsto (\Pi_x^{\mathrm{g}} \boldsymbol{f}(x))(x)$$

is itself a continuous function and

$$\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big)(x) = \big(\Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x). \tag{2.32}$$

Remark. One may wonder for which class of T-valued functions the reconstruction theorem holds. Caravenna and Zambotti proved in [23] a new notion of *germs*. A family of distributions $(f_x)_{x\in\mathbb{R}\times\mathbb{R}^d} \subset \mathcal{S}'(\mathbb{R}\times\mathbb{R}^d)$ which is measurable in x is called a germ. For any $\gamma \in \mathbb{R}$, the germ $(f_x)_x$ is called *γ-coherent* if there exists $\beta_0 \leq \gamma \wedge 0$ such that

$$\big|\big\langle f_y - f_x, p_t(x,\cdot)\big\rangle\big| \lesssim t^{\frac{\beta_0}{4}} \big(d(y,x) + t^{\frac{1}{4}}\big)^{\gamma-\beta_0}.$$

(Here, we present a simplified definition rather than the original definition in [23].) Caravenna and Zambotti [23] stated a more general reconstruction theorem at the level of coherent germs (Theorem 5.1 therein) and also stated that the coherence is actually necessary for the existence of the reconstruction (Theorem 6.1 therein). For any models $\mathsf{M} = (\Pi, g)$ and modelled distributions $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$, the x-dependent distributions $\Pi_x^{\mathrm{g}} \boldsymbol{f}(x)$ define a coherent germ.

Function-like comodules. Throughout we will work with regularity structures satisfying the following assumption saying that T and T^+ contain the polynomial regularity structure.

Assumption A1. The concrete regularity structure $((T^+, \Delta^+), (T, \Delta))$ contains the polynomial regularity structure in the following sense. One has $\mathcal{B}_\alpha^{(+)} = \{X_{(+)}^n;\, |n|_{\mathfrak{s}} = \alpha\}$, for a symbol $X_{(+)} \in T^{(+)}$ and any integer $\alpha \in \mathbb{N}$, and

$$\Delta^{(+)} X_{(+)}^n = \sum_{\ell \leq n} \binom{n}{\ell} X_{(+)}^\ell \otimes X_+^{n-\ell}.$$

The notation X_+^n allows to distinguish the elements in $\mathcal{B}^+$ and $\mathcal{B}$. Note that we always have X_+ on the right-hand side of the above tensor product while we have X or X_+ on the left-hand side depending on whether we work on T or T^+. Set

$$\mathcal{B}_X^+ := \{X_+^n;\, n \in \mathbb{N} \times \mathbb{N}^d\}, \quad \mathcal{B}_X := \{X^n;\, n \in \mathbb{N} \times \mathbb{N}^d\},$$

and write

$$\mathbf{1}_+ = X_+^0, \quad \mathbf{1} = X^0.$$

By assumption, for any *integer* α, the basis $\mathcal{B}_\alpha^{(+)}$ consists *exactly* of the elements $X_{(+)}^n$ with $|n|_{\mathfrak{s}} = \alpha$. Therefore, these polynomials are the only elements of $T^{(+)}$ with integer homogeneities. In particular, $T_0^{(+)}$ is a one-dimensional vector space.

Note the use of X_+ in the formula for ΔX^n. The space

$$T_X^+ := \mathrm{span}(\mathcal{B}_X^+)$$

with Δ^+ is isomorphic to a polynomial regularity structure, while the space

$$T_X := \mathrm{span}(\mathcal{B}_X)$$

with Δ is a right comodule over T_X^+. One defines a *canonical model* over the polynomial regularity structure

$$\mathcal{T}_X := \big((T_X^+, \Delta^+), (T_X, \Delta)\big)$$

setting for all $x, y \in \mathbb{R} \times \mathbb{R}^d$,

$$\mathrm{g}_x(X_+^n) := x^n, \quad (\Pi X^n)(y) := y^n.$$

We see that $\mathrm{g}_{yx}(X_+^n) = (y-x)^n$ and $(\Pi_x^{\mathrm{g}} X^n)(y) = (y-x)^n$, so (g, Π) is indeed a model over $\mathcal{T}_X$.

Assumption A2. Under Assumption A1, we only consider models (Π, g) whose restriction to $\mathcal{T}_X$ is the canonical model.

We only work from now on with regularity structures satisfying Assumptions A1 and A2. It is useful, to deal with sub-regularity structures of a given regularity structure, to introduce the following notion. A linear subspace V of T is called a *subcomodule* if

$$\Delta V \subset V \otimes T^+;$$

that is, defining $V_\alpha := V \cap T_\alpha$, the pair $((T^+, \Delta^+), (V, \Delta))$ is a regularity structure. A subcomodule V is said to be *function-like*, if V satisfies Assumptions A1 and A2 and if $V_\beta = 0$ whenever $\beta < 0$. Given a subcomodule V, set

$$\alpha_0(V) := \min\{\alpha \in A;\ V_\alpha \neq 0,\ \alpha \notin \mathbb{N}\}$$

if there is $\alpha \notin \mathbb{N}$ such that $V_\alpha \neq 0$; otherwise, $\alpha_0(V) := \infty$.

Corollary 7. *Let V be a function-like comodule. For a positive regularity exponent γ and $\boldsymbol{f} \in \mathcal{D}^\gamma(V, \mathrm{g})$, one has $\mathbf{R}^{\mathsf{M}} \boldsymbol{f} \in \mathcal{C}^{\alpha_0(V) \wedge \gamma}(\mathbb{R} \times \mathbb{R}^d)$ and for all $x \in \mathbb{R} \times \mathbb{R}^d$*

$$\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big)(x) = f_{\mathbf{1}}(x).$$

Proof. Set $\beta = \alpha_0(V) \wedge \gamma$. To see the regularity of f_1, we write

$$\boldsymbol{f}(x) = \sum_{|n|_{\mathfrak{s}}<\beta} \frac{f_n(x)}{n!} X^n + \sum_{\tau \in \mathcal{B}, |\tau| \geq \beta} f_\tau(x)\tau, \quad \boldsymbol{g}(x) := \sum_{\tau \in \mathcal{B}, |\tau| \geq \beta} f_\tau(x)\tau.$$

By expanding $(\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}} \boldsymbol{f}(x))_{X^n} = O(d(x,y)^{\gamma-|n|_{\mathfrak{s}}})$ for any $|n|_{\mathfrak{s}} < \beta$, we have

$$f_n(y) - \sum_{|\ell|_{\mathfrak{s}} < \gamma - |n|_{\mathfrak{s}}} \frac{f_{n+\ell}(x)}{\ell!}(y-x)^\ell - \ell!\big(\widehat{\mathrm{g}_{yx}} \boldsymbol{g}(x)\big)_{X^n} = O\big(d(x,y)^{\gamma-|n|_{\mathfrak{s}}}\big).$$

Since $(\widehat{\mathrm{g}_{yx}}\, \boldsymbol{g}(x))_{X^n} = O(d(x,y)^{\beta-|n|_{\mathfrak{s}}})$, we have

$$f_n(y) - \sum_{|\ell|_{\mathfrak{s}} < \gamma - |n|_{\mathfrak{s}}} \frac{f_{n+\ell}(x)}{\ell!}(y-x)^\ell = O\big(d(x,y)^{\beta-|n|_{\mathfrak{s}}}\big).$$

This implies that $f_0 = f_1 \in \mathcal{C}^\beta$ and $f_n = \partial^n f_1$. Therefore,

$$\begin{aligned}\big\langle f_1 - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x), p_t(x,\cdot)\big\rangle &= \left\langle f_1 - \sum_{|n|_{\mathfrak{s}}<\beta} \frac{f_n(x)}{n!}(\cdot - x)^n + \Pi_x^{\mathrm{g}} \boldsymbol{g}(x), p_t(x,\cdot)\right\rangle \\ &= \big\langle O\big(d(x,\cdot)^\beta\big) + \Pi_x^{\mathrm{g}} \boldsymbol{g}(x), p_t(x,\cdot)\big\rangle = O(t^{\frac{\beta}{4}}).\end{aligned}$$

The uniqueness part of the proof of the reconstruction theorem, Theorem 4, makes it clear that the reconstruction $\mathbf{R}^{\mathsf{M}} \boldsymbol{f}$ of $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$, with $\gamma > 0$, is characterized by the estimate

$$\big|\big\langle \mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x), p_t(x,\cdot)\big\rangle\big| \lesssim t^{\gamma'},$$

whatever positive exponent γ' appears in the upper bound. Hence, $\mathbf{R}^{\mathsf{M}} \boldsymbol{f} = f_1$. ∎

2.4. Products and derivatives

Other regularity structures than the polynomial regularity structure can be used to "model" functions. In good cases, they come equipped with a bilinear operation that plays the role played by multiplication in the usual setting and allows to define the image of a modelled distribution by a nonlinear map. This is what this section is about.

Let V, W be subcomodules of T and set

$$V_\alpha := V \cap T_\alpha, \quad W_\alpha := W \cap T_\alpha.$$

Definition. A *product* on $V \times W$ is a continuous bilinear map $\star : V \times W \to T$ such that $V_\alpha \star W_\beta \subset T_{\alpha+\beta}$ for all $\alpha, \beta \in A$. The product is said to be *regular* if

$$\Delta(\tau \star \sigma) = (\Delta\tau)(\Delta\sigma)$$

for all $\tau \in V$ and $\sigma \in W$. On the right-hand side, the product $(V \otimes T^+) \times (W \otimes T^+) \to T \otimes T^+$ is canonically defined from $\star$ and the product of T^+, setting

$$(\tau \otimes \mu)(\sigma \otimes \nu) := (\tau \star \sigma) \otimes (\mu\nu).$$

The regularity structures used in the study of singular PDEs have elements that are decorated rooted trees. The product is given as a tree product in that setting, and such a product is regular in the above sense. The details will be found in Section 9. A regular product $\star$ satisfies

$$\hat{g}(\tau \star \sigma) = \hat{g}(\tau) \star \hat{g}(\sigma)$$

for any character g on T^+. For regularity structures containing the polynomial regularity structure, one asks the following consistency assumption.

Assumption A3. Under Assumption A1, the product between T_X and T is always defined and satisfies

$$\mathbf{1} \star \tau = \tau \star \mathbf{1} = \tau \quad \text{for all } \tau \in T, \qquad X^k \star X^\ell = X^{k+\ell} \quad \text{for all } k, \ell \in \mathbb{N} \times \mathbb{N}^d.$$

We remark that Assumption A3 is not contained in Hairer's general definition [53] (of course always assumed in the specific regularity structure of decorated rooted trees [17,20, 26]). We make this assumption here because it is used in the proof of "Whitney's extension theorem" (Theorem 52 in appendix) by Martin [74]. Assumptions A1, A2, and A3 are jointly called Assumption A. The proof of the next statement is elementary and left to the reader. See the proof of [53, Theorem 4.6] if needed. For $\alpha \le 0 < \gamma$, denote by $\mathcal{D}_\alpha^\gamma$ the space of modelled distributions of the form

$$\boldsymbol{f} = \sum_{\alpha \le |\tau| < \gamma} f_\tau \tau$$

and write

$$\mathcal{Q}_{<\gamma} : T \to T_{<\gamma}$$

for the canonical projection.

Proposition 8. *Let $\alpha_1, \alpha_2 \le 0 < \gamma_1, \gamma_2$, and set $\gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)$. Let $\star : V \times W \to T$ be a regular product. Given $\boldsymbol{f}_1 \in \mathcal{D}_{\alpha_1}^{\gamma_1}(V, \mathrm{g})$ and $\boldsymbol{f}_2 \in \mathcal{D}_{\alpha_2}^{\gamma_2}(W, \mathrm{g})$, one has*

$$\mathcal{Q}_{<\gamma}(\boldsymbol{f}_1 \star \boldsymbol{f}_2) \in \mathcal{D}_{\alpha_1+\alpha_2}^{\gamma}(T, \mathrm{g}).$$

The mapping $(\boldsymbol{f}_1, \boldsymbol{f}_2) \mapsto \mathcal{Q}_{<\gamma}(\boldsymbol{f}_1 \star \boldsymbol{f}_2)$ is continuous.

Let V be a *function-like comodule* of T equipped with an associative product

$$\star : V \times V \to V.$$

Then, $\star$ is naturally extended to the multilinear map from V^n to V for any $n \ge 1$. For any $\boldsymbol{f} \in \mathcal{D}^\gamma(V, \mathrm{g})$ with $\gamma > 0$ and a smooth function $F : \mathbb{R} \to \mathbb{R}$, we define

$$F^\star(\boldsymbol{f}) := \mathcal{Q}_{<\gamma}\left(\sum_{n=0}^{\infty} \frac{F^{(n)}(f_1)}{n!} \bar{\boldsymbol{f}}^{\star n}\right), \qquad \bar{\boldsymbol{f}} := \boldsymbol{f} - f_1 \mathbf{1}.$$

The sum contains only finitely many terms since the sector V is function-like. Indeed, since $\bar{\boldsymbol{f}} = \sum_{\alpha\leq|\tau|<\gamma} f_\tau \tau$ for an $\alpha > 0$, we have $\bar{\boldsymbol{f}}^{\star n} = \sum_{n\alpha\leq|\tau|<n\gamma} g_\tau \tau$ for some coefficients g_τ. The proof of the next proposition is elementary and left to the reader; see [53, Theorem 4.15] for a proof.

Proposition 9. *Pick a positive regularity exponent γ. For any $\boldsymbol{f} \in \mathcal{D}^\gamma(V, \mathsf{g})$ and a smooth function F, one has $F^\star(\boldsymbol{f}) \in \mathcal{D}^\gamma(V, \mathsf{g})$. Moreover, the mapping $\boldsymbol{f} \mapsto F^\star(\boldsymbol{f})$ is locally Lipschitz continuous.*

Finally, we introduce a linear operator playing the role of the derivative.

Definition. A *derivative* is a continuous linear map $D : T \to T$ such that $DT_\alpha \subset T_{\alpha-1}$ for all $\alpha \in A$, and

$$\Delta(D\tau) = (D \otimes \mathrm{Id})\Delta\tau$$

for any $\tau \in T$ – by an abuse of notation, we mean $T_{\alpha-1} = \{0\}$ if $\alpha - 1 \notin A$.

The assumption on D implies

$$\hat{g}(D\tau) = D\hat{g}(\tau)$$

for any character g on T^+. From this property, it is straightforward to show the following statement.

Proposition 10. *The mapping $\mathcal{D}^\gamma(T, \mathsf{g}) \ni \boldsymbol{f} \mapsto D\boldsymbol{f} \in \mathcal{D}^{\gamma-1}(T, \mathsf{g})$ is continuous. Moreover, if $\Pi \circ D = \mathfrak{D} \circ \Pi$ holds for a first-order differential operator $\mathfrak{D}$, then*

$$\mathbf{R}^{\mathsf{M}}(D\boldsymbol{f}) = \mathfrak{D}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big)$$

for any $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathsf{g})$ with $\gamma > 1$.

3. Regularity structures built from integration operators

We describe in this section a setting where one can lift a given singular PDE into an equation set on a space of modelled distributions. The lift depends on the arbitrary choice of a model on the regularity structure, and we will see in Section 4 that the lifted equation has a unique local in time solution $\boldsymbol{u} = \boldsymbol{u}(\mathsf{M})$ for every model. The solution of the initial singular equation will be defined as the reconstruction of this unique $\boldsymbol{u}$.

The regularity structures used for the study of singular stochastic PDEs have a particular structure that comes from the fixed-point formulation of the (system of) PDE(s) under study. We concentrate here on the case where only one second-order differential operator is involved, typically $\partial_t - \Delta_x$. (See Section 9 and Appendix D for comments on the general case.) We work then with regularity structures equipped with an operator $\mathcal{I}$ that plays the role of the convolution operator $(\partial_t - \Delta_x)^{-1}$, involved in the fixed-point formulation of the equation under study. This operator is called an abstract integration map; it is introduced in Section 3.2. One associates in Section 3.3 to an abstract integration operator $\mathcal{I}$

a notion of admissible Π-maps, which roughly means that $\Pi(\mathcal{I}\tau) = (\partial_t - \Delta_x)^{-1}(\Pi\tau)$. For some models $\mathsf{M} = (g, \Pi)$ with Π admissible, the map $\mathcal{I}$ is (essentially) intertwined to the operator $(\partial_t - \Delta_x)^{-1}$ via Π. These particular models play a crucial role in Section 3.4. We construct therein a model-dependent operator $\mathcal{K}^{\mathsf{M}}$ that is (essentially) intertwined to $(\partial_t - \Delta_x)^{-1}$ via the reconstruction map R^{M}, and which also has a regularizing property analogous to a similar property enjoyed by $(\partial_t - \Delta_x)^{-1}$. The operator $\mathcal{K}^{\mathsf{M}}$ will be used in Section 4 to lift the singular PDE into an equation on a space of modelled distributions. The condition that Π is admissible has some far-reaching consequences, and it is not obvious in the first place that one can construct a non-trivial model that is admissible. Section 3.5 is dedicated to constructing a large class of admissible models.

The main result of this section is Theorem 17, which gives the continuity property of the operator $\mathcal{K}^{\mathsf{M}}$.

A remark is in order before we set the scene in Section 3.1. We will restrict our study to regularity structures for which the minimum homogeneity of its elements satisfies

$$\beta_0 = \min A > -2. \tag{3.1}$$

This condition ensures that elements of the form $\mathcal{I}(\tau)$ that appear in the expansion of solutions to the regularity structure lift of the considered class of singular stochastic PDE are of positive homogeneity. (So, if $\mathcal{I}$ were increasing the homogeneity of all symbols by a, we would require $\beta_0 > -a$.) While the generalized (KPZ) equation (1.3) satisfies for instance this assumption, not all singular stochastic PDEs satisfy it. This is, for instance, the case of the Φ_2^4, Φ_3^4 or sine-Gordon equations. This kind of equations can nonetheless be studied within the setting of regularity structures by writing their solutions as the sum of an explicit functional of the noise and a remainder term that solves an equation that can be formulated in a regularity structure satisfying condition (3.1) – the so-called da Prato–Debussche trick, after similar operation was used in their work [31]. We consider as an example the case of the Φ_3^4 equation

$$(\partial_t - \Delta_x)u = -u^3 + \zeta,$$

set on the 3-dimensional torus, with ζ a spacetime white noise of Hölder regularity $(-5/2)^-$. We decompose a priori the solution u into $u = X + v$, where

$$\begin{aligned}(\partial_t - \Delta_x)X &= \zeta,\\ (\partial_t - \Delta_x)v &= -v^3 - 3v^2X - 3vX^2 - X^3.\end{aligned}$$

The polynomial functions $(X^n)_{1\le n\le 3}$ can directly be defined as elements of $\mathcal{C}^{(-n/2)^-}$ by probabilistic means. The above equation for v can then be formulated in a regularity structure with three noise symbols for X, X^2, and X^3, in which $\beta_0 = (-3/2)^- > -2$. The interested reader will find more details on this matter for a general class of singular stochastic PDEs in Section 5 of Bruned, Chandra, Chevyrev, and Hairer's work [17].

3.1. Operators on $\mathbb{R} \times \mathbb{R}^d$

We will be interested in (systems of) singular stochastic PDEs that involve possibly two types of differential operators: the derivatives ∂_i in the directions of the canonical basis of $\mathbb{R} \times \mathbb{R}^d$ and the second-order differential operator

$$\mathbf{L} := \partial_{x_0} - \Delta_{x'} + 1.$$

Denote by $\mathbf{L}^{-1}$ the resolution operator associating to a Schwartz function $v \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ the solution $u \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ to the equation

$$\mathbf{L}u = v.$$

The strict positiveness of $-\Delta_{x'} + 1$ ensures uniqueness of a solution $u \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ to the preceding equation. (This is the reason why we work with $\mathbf{L}$ rather than with the heat operator.) The operator $\mathbf{L}^{-1}$ can be represented using a variant of the elliptic operators $\mathcal{G}$ introduced in Section 2.3. Indeed, we have

$$\begin{aligned}\big(\partial_{x_0} - \Delta_{x'} + 1\big)^{-1} &= -\big(\partial_{x_0} + \Delta_{x'} - 1\big)\big(-\partial_{x_0}^2 + (\Delta_{x'} - 1)^2\big)^{-1} \\ &= -\int_0^\infty (\partial_{x_0} + \Delta_{x'} - 1) e^{r(\partial_{x_0}^2 - (\Delta_{x'}-1)^2)} dr.\end{aligned}$$

Write

$$\mathbf{L}^{-1} = -\int_0^\infty (\partial_{x_0} + \Delta_{x'} - 1) e^{r(\partial_{x_0}^2 - (\Delta_{x'}-1)^2)} dr =: \int_0^\infty K_r \, dr.$$

We thus use the inhomogeneous operator

$$\widetilde{\mathcal{G}} := \partial_{x_0}^2 - (\Delta_{x'} - 1)^2$$

instead of the operator $\mathcal{G}$ considered in Section 2.3. The contents in the previous section hold similarly even if we redefine p_t as the kernel of $e^{t\widetilde{\mathcal{G}}}$. Since

$$q_t(x, y) = -(\partial_{x_0} + \Delta_{x'} - 1) p_t(x, y),$$

by Proposition 2, the kernel $q_r(x, y)$ of the operator K_r satisfies the x-uniform bounds

$$\big|\big\langle \Pi_x^{\mathsf{g}} \tau, \partial_x^n q_r(x, \cdot)\big\rangle\big| \lesssim r^{\frac{|\tau| - |n|_{\mathfrak{s}} - 2}{4}} \tag{3.2}$$

for any $\tau \in \mathcal{B}$, $r \in (0, 1]$, and $n \in \mathbb{N} \times \mathbb{N}^d$, with the exponent $|n|_{\mathfrak{s}} + 2$ coming from the derivative operators ∂_x^n and $(\partial_{x_0} + \Delta_{x'} - 1)$ applied to $e^{r\widetilde{\mathcal{G}}}$. It is convenient, for technical purposes, we replace q_r with the function

$$\tilde{q}_r(x, y) := q_r(x, y) - P_r(\partial_x) q_1(x, y)$$

for some r-dependent polynomial

$$P_r(\xi) = P_r(\xi_0, \xi_1, \dots, \xi_d)$$

whose coefficients are bounded over $r \in [0, 1]$, chosen to satisfy the property that

$$\int_{\mathbb{R}\times\mathbb{R}^d} y^n \tilde{q}_r(x, y)dy = 0 \tag{3.3}$$

for any $n \in \mathbb{N} \times \mathbb{N}^d$ such that $|n|_{\mathfrak{s}} < N$ for fixed positive number N. For instance, when $N = 2$, we choose

$$\tilde{q}_r(x, y) := q_r(x, y) - e^{-(r-1)} q_1(x, y).$$

In general, by noting that Fourier transform of $Q_r(\cdot) := q_r(\cdot, 0)$ is given by $\widehat{Q_r}(\xi) = f(\xi)e^{rg(\xi)}$ for some polynomials f and g, and that

$$Q_r(\xi) - P_r(\xi)Q_1(\xi) = Q_1(\xi)\big(e^{(r-1)g(\xi)} - P_r(\xi)\big),$$

we can choose a polynomial $P_r(\xi)$ such that $\partial_\xi^n(e^{(r-1)g(\xi)} - P_r(\xi))|_{\xi=0}$ for any $|n|_{\mathfrak{s}} < N$. By using this modified kernel, we decompose $\mathbf{L}^{-1}$ under the form

$$\mathbf{L}^{-1} = \mathbf{K} + \mathbf{K}',$$

with

$$\mathbf{K} := \int_0^1 K_r\, dr - \left(\int_0^1 P_r(\partial)K_1 dr\right), \quad \mathbf{K}' := \int_1^\infty K_r\, dr + \left(\int_0^1 P_r(\partial)K_1 dr\right).$$

It is elementary to see that the operators $\mathbf{K}'$ map $\mathcal{C}^\gamma(\mathbb{R} \times \mathbb{R}^d)$ into $\mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^d)$ for any regularity exponent $\gamma \in \mathbb{R}$. This decomposition is similar but different to the one $\bar{K} = K + R$ as in [53, Lemma 5.5], where $\bar{K}$ is the Green function, K is a singular kernel with a bounded support, and R is a smooth remainder. We concentrate on the operator $\mathbf{K}$ in the remainder of this section. Denote by

$$K(x, y) := \int_0^1 \tilde{q}_r(x, y)dr \tag{3.4}$$

its kernel. The compensation of K ensures that

$$\int_{\mathbb{R}\times\mathbb{R}^d} y^n K(x, y)dy = 0$$

for any $|n|_{\mathfrak{s}} < N$ with fixed positive number N. This property is used in the proof of Theorem 17 in this section and the proof of Theorem 44 later.

Proposition 11 (Schauder estimates for $\mathbf{L}^{-1}$). *The operators $\mathbf{L}^{-1}$ and $\mathbf{K}$ are continuous operators from $\mathcal{C}^\gamma(\mathbb{R} \times \mathbb{R}^d)$ into $\mathcal{C}^{\gamma+2}(\mathbb{R} \times \mathbb{R}^d)$ for all non-integer regularity exponents $\gamma \in \mathbb{R}$.*

Proof. It is sufficient to show the estimate for $\widetilde{\mathbf{K}} = \int_0^1 K_r\, dr$. Note that

$$K_t = -e^{t\widetilde{\mathscr{G}}}(\partial_{x_0} + \Delta_{x'} - 1),$$

as $\widetilde{\mathscr{G}}$ and $(\partial_{x_0} + \Delta_{x'} - 1)$ commute. We use the freedom on the choice of the $((2, 1, \dots, 1)$-scaling) elliptic operator used to define the Hölder spaces, while giving equivalent norms, to work with the norm associated with the operator $\widetilde{\mathscr{G}}$ rather than the operator $\mathscr{G}$. We emphasize that fact by writing $\widetilde{\mathcal{Q}}_s^{(N)}$ for the operators built from $\widetilde{\mathscr{G}}$ in the same way as $\mathcal{Q}_s^{(N)}$ is built from $\mathscr{G}$. Given a distribution $\Lambda \in \mathcal{C}^\gamma(\mathbb{R} \times \mathbb{R}^d)$, with $\gamma \in \mathbb{R}$ non-integer, we read on the identity

$$\widetilde{\mathcal{Q}}_s^{(N)}\big(\widetilde{\mathbf{K}}(\Lambda)\big) = \int_0^1 \widetilde{\mathcal{Q}}_s^{(N)}\big(K_r(\Lambda)\big)dr = -\int_0^1 \left(\frac{s}{s+r}\right)^N \widetilde{\mathcal{Q}}_{s+r}^{(N)}\big((\partial_{x_0} + \Delta_{x'} - 1)\Lambda\big)dr$$

the estimate

$$\big\| Q_s^{(N)}\big(\widetilde{\mathbf{K}}(\Lambda)\big)\big\|_\infty \lesssim s^{\frac{\gamma-2}{4}+1} + O(s^N).$$

The result follows for all $\gamma + 2 < 4N$. The equivalence of the different Hölder norms corresponding to different choices of N gives the conclusion. ■

(We refer the reader to Section 14.3 of the second edition of Friz and Hairer's lecture notes [43] for a particularly nice proof of the classical Schauder estimates using different tools.) For a regularity structure $\mathcal{T}$ for which $\beta_0 = \min A > -2$, and a model (Π, g) on it, Schauder estimates imply in particular that all the distributions $\mathbf{K}(\Pi\tau)(\cdot)$, hence all the distributions $\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)$, are actually defined pointwise for any $x \in \mathbb{R} \times \mathbb{R}^d$, making sense of $\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x)$, or even $\partial^n\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x)$, for $|n|_{\mathfrak{s}} < \beta_0 + 2$. The following lemma allows to take profit from the fact that $\Pi_x^{\mathsf{g}}\tau$ behaves near x "as" an element of $\mathcal{C}^{|\tau|}(\mathbb{R} \times \mathbb{R}^d)$, to give meaning to $\partial^n\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x)$, for all multi-indices n such that $|n|_{\mathfrak{s}} < |\tau| + 2$.

Lemma 12. *Assume $\beta_0 > -2$. Given $\tau \in \mathcal{B}$ and $n \in \mathbb{N} \times \mathbb{N}^d$, the integral*

$$\big(\partial^n\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)\big)(x) := \big\langle \Pi_x^{\mathsf{g}}\tau, \partial_x^n K(x, \cdot)\big\rangle := \int_0^1 \big\langle \Pi_x^{\mathsf{g}}\tau, \partial_x^n \widetilde{q}_r(x, \cdot)\big\rangle dr \tag{3.5}$$

converges for all $x \in \mathbb{R} \times \mathbb{R}^d$, provided $|n|_{\mathfrak{s}} < |\tau| + 2$.

Proof. It follows from (3.2) that the first term on the right-hand side of (3.5) is integrable over $r \in (0, 1)$ if

$$|\tau| - |n|_{\mathfrak{s}} > -2.$$ ■

The $\mathbb{R} \times \mathbb{R}^d$-indexed distributions $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x)$ satisfy a similar bound to (3.2) for any modelled distribution $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathsf{g})$. We can then define properly $\partial^n\mathbf{K}(\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x))(x)$ for all multi-indices n such that $|n|_{\mathfrak{s}} < \gamma + 2$, as in the preceding lemma.

3.2. Regularity structures with some abstract integration operators

Recall Assumption A essentially says that we consider regularity structures containing the canonical polynomial structure and models that behave naturally on the latter. In this section, we consider a regularity structure that can represent more specific functions/distributions. To explain the motivation of the next assumption, pick a γ-Hölder function f with some $\gamma > 0$ and let us try to add a new symbol F with homogeneity γ which represents f to a polynomial regularity structure $\mathcal{T}_X$. The new model space is $T = \mathrm{span}(\mathcal{B}_X \cup \{F\})$. On the other hand, T^+ is required to be rich enough to define the application of the model to F as

$$\Pi F = f, \quad (\Pi_x^{\mathrm{g}} F)(\cdot) = f(\cdot) - \sum_{|n|_{\mathfrak{s}}<\gamma} \frac{(\cdot - x)^n}{n!} \partial^n f(x) = O\big(d(\cdot, x)^\gamma\big). \tag{3.6}$$

To represent coefficients $\partial^n f$, it would be natural to introduce symbols $\{F_n\}_{|n|_{\mathfrak{s}}<\gamma}$ in T^+. Then, a good definition of Δ is given by

$$\Delta F = F \otimes \mathbf{1} + \sum_{|n|_{\mathfrak{s}}<\gamma} \frac{X^n}{n!} \otimes F_n.$$

Indeed, by applying $\Pi_x^{\mathrm{g}} \otimes \mathrm{g}_x$ on both sides and noting that $\Pi = (\Pi_x^{\mathrm{g}} \otimes \mathrm{g}_x)\Delta$, we have the second equality of (3.6) under the choice $\mathrm{g}_x(F_n) = \partial^n f(x)$. Similarly, we also need to define the application of Δ^+ to each F_n by

$$\Delta^+ F_n = F_n \otimes \mathbf{1} + \sum_{|m|_{\mathfrak{s}}<\gamma-|n|_{\mathfrak{s}}} \frac{X^m}{m!} \otimes F_{n+m}$$

to define $\mathrm{g}_{yx}(F_n)$ as a remainder term of the Taylor expansion of $\partial^n f$. For this reason, it would be natural to consider the regularity structure that satisfies in addition to Assumption A the following set of assumptions. Recall that we denote by $\{e_0, e_1, \dots, e_d\}$ the canonical basis of $\mathbb{N} \times \mathbb{N}^d$.

Assumption B1. (a) The basis $\mathcal{B}^+$ of T^+ is a commutative monoid with unit $\mathbf{1}_+$, freely generated by the symbols

$$\{X_+^{e_i}\}_{0\le i\le d} \cup \big\{\mathcal{I}_n^+ \tau\big\}_{\tau\in\mathcal{B},\, n\in\mathbb{N}\times\mathbb{N}^d,\, |\tau|+2-|n|_{\mathfrak{s}}>0},$$

where each element has homogeneity

$$\big|X_+^{e_i}\big| := \mathfrak{s}_i, \quad \big|\mathcal{I}_n^+ \tau\big| := |\tau| + 2 - |n|_{\mathfrak{s}}.$$

The operators Δ and Δ^+ are related by the *intertwining relations*

$$\Delta^+(\mathcal{I}_n^+ \tau) = (\mathcal{I}_n^+ \otimes \mathrm{Id})\Delta\tau + \sum_{\ell\in\mathbb{N}\times\mathbb{N}^d,\, |\ell|_{\mathfrak{s}}<|\tau|+2-|n|_{\mathfrak{s}}} \frac{X_+^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+ \tau \tag{3.7}$$

for any $\tau \in \mathcal{B}$.

(b) For $n \in \{0, e_1, \dots, e_d\}$ – i.e., $n \in \mathbb{N} \times \mathbb{N}^d$ such that $|n|_{\mathfrak{s}} \le 1$, there are operators $\mathcal{I}_n : T \to T$, with

$$\big|\mathcal{I}_n \tau\big| := |\tau| + 2 - |n|_{\mathfrak{s}}, \quad \tau \in \mathcal{B}.$$

One has for any $\tau \in \mathcal{B}$

$$\Delta(\mathcal{I}_n \tau) = (\mathcal{I}_n \otimes \mathrm{Id})\Delta\tau + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d,\, |\ell|_{\mathfrak{s}} < |\tau|+2-|n|_{\mathfrak{s}}} \frac{X^\ell}{\ell!} \otimes \mathcal{I}^+_{n+\ell}\tau. \tag{3.8}$$

For simplicity, we write

$$\mathcal{I}^{(+)} := \mathcal{I}^{(+)}_0.$$

The operator $\mathcal{I}_n$ is an abstract version of the convolution operator $\partial^n \mathbf{K}$. The restriction $|n|_{\mathfrak{s}} \leq 1$ on n means that we only consider $\mathbf{K}$ or $\partial_i \mathbf{K}$; this is sufficient for the study of all (systems of) singular stochastic PDEs whose solutions are functions involving second-order differential operators satisfying the above classical Schauder estimate. Two remarks are in order.

(a) The main point of Assumption B1 is the introduction of the operator $\mathcal{I}_n$. It brings the supplementary operator $\mathcal{I}^+_n$ for the reasons mentioned before. The first term on the right-hand side of the identity (3.8) seems different from the definition of ΔF mentioned before, but it specifies the action of the recentering operations on τ. Indeed, by applying $\Pi^{\mathrm{g}}_x \otimes \mathrm{g}_x$ on both sides and noting that $\widehat{\mathrm{g}_x}(\tau) = \tau + \sum_{\sigma<\tau} \mathrm{g}_x(\tau/\sigma)\sigma$, we have

$$\Pi^{\mathrm{g}}_x(\mathcal{I}_n\tau) = \Pi\mathcal{I}_n\tau - \sum_{\sigma<\tau} \mathrm{g}_x(\tau/\sigma)\Pi^{\mathrm{g}}_x\mathcal{I}_n\sigma - \sum_{\ell} \mathrm{g}_x(\mathcal{I}^+_{n+\ell}\tau)\frac{(\cdot - x)^n}{n!}.$$

(b) We assume that T^+ is entirely constructed from the $\mathcal{I}^+_n$ operators and the polynomials and has no other elements. This is because these elements are rich enough to build the regularity structure including polynomials, operators $\mathcal{I}_n$, and their (possible) products. The regularity structures that are used for the study of singular stochastic PDEs have the same structure as above, which is described in detail in Section 9.

We notice here some algebraic formulas about Δ and Δ^+. We let the readers check that identity (3.7) ensures the co-associativities

$$(\Delta^{(+)} \otimes \mathrm{Id})\Delta^{(+)}(\mathcal{I}^{(+)}_n\tau) = (\mathrm{Id} \otimes \Delta^+)\Delta^{(+)}(\mathcal{I}^{(+)}_n\tau)$$

on elements of $T^{(+)}$ of the form $\mathcal{I}^{(+)}_n\tau$. Next, using identity (3.7), we check that the antipode S_+ on T^+ satisfies the inductive relation

$$S_+(\mathcal{I}^+_n\tau) = -\sum_{\ell\in\mathbb{N}\times\mathbb{N}^d} \frac{(-X)^\ell}{\ell!}\mathcal{M}_+(\mathcal{I}^+_{n+\ell} \otimes S_+)\Delta\tau, \tag{3.9}$$

where we denote by $\mathcal{M}_+$ the multiplication operator in the algebra T^+. Together with the relation $S_+(X_+) = -X_+$, such a formula defines indeed a unique algebra morphism.

Recall from Appendix B the defining property (B.1) of the antipode S_+ on T^+. As T^+ is a Hopf algebra by Proposition 48, it suffices to see that

$$\mathcal{M}_+\big(\mathrm{Id}\otimes S_+\big)\Delta^+\big(\mathcal{I}_n^+\tau\big)=0$$

for all $n\in\mathbb{N}\times\mathbb{N}^d$ and $\tau\in T$. This relation follows from (3.9) and (3.7), writing

$$\begin{aligned}
&\mathcal{M}_+\big(\mathrm{Id}\otimes S_+\big)\Delta^+\big(\mathcal{I}_n^+\tau\big)\\
&\quad=\mathcal{M}_+\big(\mathcal{I}_n^+\otimes S_+\big)\Delta\tau+\sum_{\ell\in\mathbb{N}\times\mathbb{N}^d}\frac{X^\ell}{\ell!}S_+\big(\mathcal{I}_{n+\ell}^+\tau\big)\\
&\quad=\Big\{\mathcal{M}_+\big(\mathcal{I}_n^+\otimes S_+\big)-\sum_{\ell,k\in\mathbb{N}\times\mathbb{N}^d}\frac{X^\ell}{\ell!}\frac{(-X)^k}{k!}\mathcal{M}_+\big(\mathcal{I}_{n+\ell+k}^+\otimes S_+\big)\Big\}\Delta\tau\\
&\quad=\Big\{\mathcal{M}_+\big(\mathcal{I}_n^+\otimes S_+\big)-\sum_{\ell\in\mathbb{N}\times\mathbb{N}^d}\frac{(X-X)^\ell}{\ell!}\mathcal{M}_+\big(\mathcal{I}_{n+\ell}^+\otimes S_+\big)\Big\}\Delta\tau\\
&\quad=0.
\end{aligned}$$

Remark that the image by the operator $\mathcal{I}$ of a modelled distribution is not a modelled distribution. The next two sections are dedicated to constructing a model-dependent map $\mathcal{K}^{\mathsf{M}}$ that maps continuously all $\mathcal{D}^\gamma(T,\mathrm{g})$ into $\mathcal{D}^{\gamma+2}(T,\mathrm{g})$ when γ is a positive non-integer real number – the analogue of part of Schauder estimates – and is intertwined to the convolution operator $\mathbf{K}$

$$\mathbf{K}\circ\mathbf{R}^{\mathsf{M}}=\mathbf{R}^{\mathsf{M}}\circ\mathcal{K}^{\mathsf{M}}$$

via the reconstruction operator $\mathbf{R}^{\mathsf{M}}$ associated with M. We say that $\mathcal{K}^{\mathsf{M}}$ is a *lift of* $\mathbf{K}$. The construction of this operator requires the introduction of the notion of admissible model.

3.3. Admissible models

In this section, we consider only the operator $\mathcal{I}=\mathcal{I}_0:T\to T$. The following notion plays a key role in the proof of the existence of a lift of the convolution operator $\mathbf{K}$. Recall that $\mathbf{K}(\zeta)$ is well defined pointwise for any distribution $\zeta\in\mathcal{C}^\beta(\mathbb{R}\times\mathbb{R}^d)$ with $\beta>-2$, by Lemma 12. The following definition is directly from [20, Definition 6.9].

Definition. Let $\mathcal{T}$ be a regularity structure satisfying Assumptions A3 and B1. Assume $\beta_0>-2$. A Π-map on T is said to be **K**-*admissible* if it satisfies

$$\Pi(\mathcal{I}\tau)=\mathbf{K}(\Pi\tau),\tag{3.10}$$

and

$$\Pi(X^n\star\tau)(x)=x^n(\Pi\tau)(x)$$

for any $\tau\in\mathcal{B}$ and $n\in\mathbb{N}\times\mathbb{N}^d$. A model (Π,g) on $\mathcal{T}$ is said to be **K**-admissible if its Π-map is **K**-admissible.

The notion of admissible Π-map gives flesh to the idea that the operator $\mathcal{I}$ is the regularity structure counterpart of the convolution operator $\mathbf{K}$. The importance of the notion of $\mathbf{K}$-admissible model comes from Theorem 17 in the next section. It shows that when working with $\mathbf{K}$-admissible models M, one can upgrade the intertwining relation (3.10) into an intertwining relation between $\mathbf{K}$ and an operator $\mathcal{K}^{\mathsf{M}}$ on modelled distributions, via the reconstruction map $\mathbf{R}^{\mathsf{M}}$ associated with M.

While for a general model as in Definition 1 defining a g-map satisfying the constraint (2.19) is decorrelated from the task of defining a Π-map satisfying the constraint (2.20) it will turn out that imposing the intertwining relation (3.10) on Π will constrain strongly g. Unlike general models, admissible models on a regularity structure satisfying Assumptions A and B1 will turn out to be partly defined by their Π map; this will be proved in Proposition 15. The g-map of an admissible model on a regularity structure satisfying further a mild Assumption B2 will turn out to be *entirely* defined by its Π map.

We worked so far with models that are not constrained by anything else than their defining properties (2.19) and (2.20) and it is not clear that one can further impose additional conditions like (3.10). We will construct in Section 3.5 a whole class of admissible models with values in the set of smooth functions. This is all we need for the study of singular stochastic PDEs, as the nonsmooth admissible models involved in this setting are limits of smooth admissible models, and limits of admissible models are admissible. As for now, we keep going and see what can be done with admissible models.

Recall from Lemma 12 the definition of $\partial^n \mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x)$, for any $x \in \mathbb{R}\times\mathbb{R}^d$ and $n \in \mathbb{N}\times\mathbb{N}^d$ such that $|n|_{\mathfrak{s}} < |\tau| + 2$, and define the model-dependent polynomial-valued function on T

$$\mathcal{J}^{\mathsf{M}}(x)\tau := \sum_{|n|_{\mathfrak{s}}<|\tau|+2} \frac{X^n}{n!}\partial^n \mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x) \in T_X$$

for any $\tau \in \mathcal{B}$ and $x \in \mathbb{R}\times\mathbb{R}^d$.

Proposition 13. *For a* $\mathbf{K}$*-admissible model* M *on* $\mathcal{T}$*, one has, for any* $x \in \mathbb{R}\times\mathbb{R}^d$ *and* $\tau \in T$*,*

$$\Pi_x^{\mathsf{g}}\big(\mathcal{I}\tau + \mathcal{J}^{\mathsf{M}}(x)\tau\big) = \mathbf{K}(\Pi_x^{\mathsf{g}}\tau).$$

Proof. Using identity (3.8) and the admissibility of the model, one has indeed

$$\begin{aligned}
\Pi_x^{\mathsf{g}}(\mathcal{I}\tau) &= (\Pi\otimes \mathsf{g}_x^{-1})(\Delta\mathcal{I}\tau) \\
&= \sum_{\sigma\leq\tau}(\Pi\otimes\mathsf{g}_x^{-1})(\mathcal{I}\sigma\otimes\tau/\sigma) + \sum_{|\ell|_{\mathfrak{s}}<|\tau|+2}\frac{(\cdot)^\ell}{\ell!}\,\mathsf{g}_x^{-1}(\mathcal{I}_\ell^+\tau) \\
&=: \sum_{\sigma\leq\tau}\mathbf{K}(\Pi\sigma)\mathsf{g}_x^{-1}(\tau/\sigma) + P_x(\tau,\cdot) = \mathbf{K}\big(\Pi_x^{\mathsf{g}}\tau\big) + P_x(\tau,\cdot),
\end{aligned}$$

so $\Pi_x^{\mathsf{g}}(\mathcal{I}\tau)$ and $\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)$ differ by a polynomial $P_x(\tau,\cdot)$ of degree at most $|\tau|+2$. We identify this polynomial with $-\Pi_x^{\mathsf{g}}(\mathcal{J}^{\mathsf{M}}(x)\tau)$, noting that $\Pi_x^{\mathsf{g}}(\mathcal{I}\tau)$ has null derivatives at x up to the order the integer part of $|\tau|+2$, from condition (2.20). ∎

Corollary 14. *For a* $\mathbf{K}$*-admissible model* M *on* $\mathcal{T}$*, one has, for any* $x, y \in \mathbb{R} \times \mathbb{R}^d$*,*

$$\widehat{\mathsf{g}_{yx}}\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big) = (\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y))\widehat{\mathsf{g}_{yx}}. \tag{3.11}$$

Proof. Given $\tau \in T$ and $x, y \in \mathbb{R} \times \mathbb{R}^d$, one has both

$$(\widehat{\mathsf{g}_{yx}}(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)) - (\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y))\widehat{\mathsf{g}_{yx}})\tau \in T_X,$$

and

$$\begin{aligned}
&\Pi_y^{\mathsf{g}}\big(\widehat{\mathsf{g}_{yx}}\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\tau - \big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y)\big)\widehat{\mathsf{g}_{yx}}\tau\big) \\
&\quad = \Pi_x^{\mathsf{g}}\big(\mathcal{I}\tau + \mathcal{J}^{\mathsf{M}}(x)\tau\big) - \Pi_y^{\mathsf{g}}\big(\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y)\big)\widehat{\mathsf{g}_{yx}}\tau\big) \\
&\quad = \mathbf{K}\big(\Pi_x^{\mathsf{g}}\tau\big) - \mathbf{K}\big(\Pi_y^{\mathsf{g}}\widehat{\mathsf{g}_{yx}}\tau\big) = 0,
\end{aligned}$$

from Proposition 13. As Π_y^{g} is injective on T_X, this implies that

$$\widehat{\mathsf{g}_{yx}}\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big) - \big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y)\big)\widehat{\mathsf{g}_{yx}} = 0. \qquad \blacksquare$$

Proposition 15. *Let* $\mathcal{T}$ *be a regularity structure satisfying Assumptions* A *and* B1*. The* g*-map of a* $\mathbf{K}$*-admissible model* (Π, g) *on* $\mathcal{T}$ *satisfies*

$$\mathsf{g}_x(\mathcal{I}^+\tau) = \mathbf{K}(\Pi\tau)(x)$$

for any $\tau \in \mathcal{B}$ *and all* $x \in \mathbb{R} \times \mathbb{R}^d$*.*

Proof. We show that there is at most one choice of $\mathsf{g}(\mathcal{I}^+\tau)$ such that (Π, g) is an admissible model. Applying $\Pi \otimes \mathsf{g}_x^{-1}$ to the identity (3.8) giving $\Delta(\mathcal{I}_n\tau)$, with $n = 0$, one gets from the $\mathbf{K}$-admissibility of Π

$$\begin{aligned}
\Pi_x^{\mathsf{g}}(\mathcal{I}\tau) &= \mathbf{K}\big(\Pi_x^{\mathsf{g}}\tau\big) + \sum_{|\ell|_{\mathfrak{s}} < |\tau|+2} \frac{(\cdot)^\ell}{\ell!}\, \mathsf{g}_x^{-1}(\mathcal{I}_\ell^+\tau) \\
&= \mathbf{K}\big(\Pi_x^{\mathsf{g}}\tau\big) + \sum_{|\ell|_{\mathfrak{s}} < |\tau|+2} \frac{(\cdot - x)^\ell}{\ell!}\, \mathsf{f}_x(\mathcal{I}_\ell^+\tau),
\end{aligned}$$

where f and g are related by the formulas

$$\begin{aligned}
\mathsf{f}_x(\mathcal{I}_\ell^+\tau) &:= \sum_{|m|_{\mathfrak{s}} < |\tau|+2-|\ell|_{\mathfrak{s}}} \frac{x^m}{m!}\mathsf{g}_x^{-1}(\mathcal{I}_{\ell+m}^+\tau), \\
\mathsf{g}_x^{-1}(\mathcal{I}_n^+\tau) &= \sum_{|m|_{\mathfrak{s}} < |\tau|+2-|n|_{\mathfrak{s}}} \frac{(-x)^m}{m!}\mathsf{f}_x(\mathcal{I}_{n+m}^+\tau).
\end{aligned}$$

As in the proof of Proposition 13, since the derivatives of $\Pi_x^{\mathsf{g}}(\mathcal{I}\tau)$ up to order $|\tau| + 2$ vanish at x, we have

$$\mathsf{f}_x(\mathcal{I}_n^+\tau) = -\partial^n\mathbf{K}(\Pi_x^{\mathsf{g}}\tau)(x);$$

hence,

$$\mathrm{g}_x^{-1}(\mathcal{I}_n^+\tau) = -\sum_{|m|_{\mathfrak{s}}<|\tau|+2-|n|_{\mathfrak{s}}} \frac{(-x)^m}{m!}\partial^{n+m}\mathbf{K}\big(\Pi_x^{\mathrm{g}}\tau\big)(x). \tag{3.12}$$

This implies another inductive formula

$$\mathrm{g}_x(\mathcal{I}_k^+\tau) = \sum_{\sigma\leq\tau;|k|_{\mathfrak{s}}<|\sigma|+2} \mathrm{g}_x(\tau/\sigma)\partial^k\mathbf{K}(\Pi_x^{\mathrm{g}}\sigma)(x), \tag{3.13}$$

which is proved by applying $\mathrm{g}_x^{-1}\otimes\mathrm{g}_x$ to the identity (3.7) describing $\Delta^+(\mathcal{I}_k^+\tau)$ and using (3.12). Since $\beta_0 > -2$, if $k = 0$, the condition $|k|_{\mathfrak{s}} < |\sigma| + 2$ can be removed. Hence, we have $\mathrm{g}_x(\mathcal{I}^+\tau) = \mathbf{K}(\Pi\tau)(x)$, by the comodule identity (2.8). ∎

Set

$$\mathcal{J}^{\mathsf{M}+}(x)\tau := \sum_{|n|_{\mathfrak{s}}<|\tau|+2} \frac{X_+^n}{n!}\partial^n\mathbf{K}(\Pi_x^{\mathrm{g}}\tau)(x) \in T_X^+ \subset T^+$$

for any $\tau\in\mathcal{B}$ and $x\in\mathbb{R}\times\mathbb{R}^d$. The following statement is proved exactly as Proposition 13 and Corollary 14; it will be used in the proof of Theorem 19.

Proposition 16. *Given a regularity structure satisfying Assumptions* A *and* B1 *and a* **K**-*admissible model* (Π, g) *on it, one has, for any* $x, y \in \mathbb{R}\times\mathbb{R}^d$,

$$\mathrm{g}_{yx}\big(\mathcal{I}^+\tau + \mathcal{J}^{\mathsf{M}+}(x)\tau\big) = \mathbf{K}(\Pi_x^{\mathrm{g}}\tau)(y)$$

and

$$\widehat{\mathrm{g}_{yx}}^+\big(\mathcal{I}^+ + \mathcal{J}^{\mathsf{M}+}(x)\big) = \big(\mathcal{I}^+ + \mathcal{J}^{\mathsf{M}+}(y)\big)\widehat{\mathrm{g}_{yx}}.$$

Recall here that $\hat{g}^+ := (\mathrm{Id}\otimes g)\Delta^+ : T^+\to T^+$ *denotes the action of* $g\in G^+$ *on* T^+ *defined in the same way as* $\hat{g}: T\to T$.

3.4. Lifting K as a continuous map from $\mathcal{D}^\gamma(T, \mathrm{g})$ into $\mathcal{D}^{\gamma+2}(T, \mathrm{g})$

For a given **K**-admissible model $\mathsf{M} = (\Pi, \mathrm{g})$, we define in this section a continuous map $\mathcal{K}^{\mathsf{M}}$ from $\mathcal{D}^\gamma(T, \mathrm{g})$ into $\mathcal{D}^{\gamma+2}(T, \mathrm{g})$, for any positive non-integer regularity exponent γ, intertwined to **K** via the reconstruction operator

$$\mathbf{K}\circ\mathbf{R}^{\mathsf{M}} = \mathbf{R}^{\mathsf{M}}\circ\mathcal{K}^{\mathsf{M}}. \tag{3.14}$$

To get a grasp on what $\mathcal{K}^{\mathsf{M}}$ could be, one keeps from the reconstruction theorem, Theorem 4, the image that, for $\boldsymbol{g}\in\mathcal{D}^{\gamma+2}(T, \mathrm{g})$ and $x\in\mathbb{R}\times\mathbb{R}^d$, the distribution $\mathbf{R}^{\mathsf{M}}\boldsymbol{g} - \Pi_x^{\mathrm{g}}\boldsymbol{g}(x)$ behaves near x like the function $d(\cdot, x)^{\gamma+2}$. For $\boldsymbol{f}\in\mathcal{D}^\gamma(T, \mathrm{g})$, since we have

$$\mathbf{K}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big) - \Pi_x^{\mathrm{g}}\big(\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\boldsymbol{f}(x)\big) = \mathbf{K}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x)\big),$$

from Proposition 13, it then looks natural to add to $(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x))\boldsymbol{f}(x)$ the polynomial expansion

$$\big(\mathcal{N}^{\mathsf{M}}\boldsymbol{f}\big)(x) := \sum_{|\ell|_{\mathfrak{s}}<\gamma+2} \frac{X^\ell}{\ell!}\big(\partial^\ell\mathbf{K}\big)\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x)\big)(x)$$

of $\mathbf{K}(\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x))$ at point x, at order $\gamma+2$, and expect that

$$\mathbf{K}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big) - \Pi_x^{\mathsf{g}}\Big(\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\boldsymbol{f}(x) + \big(\mathcal{N}^{\mathsf{M}}\boldsymbol{f}\big)(x)\Big)$$

behaves like $d(\cdot,x)^{\gamma+2}$ near x. (The remark after Lemma 12 justifies the good definition of the quantities $(\partial^\ell\mathbf{K})(\mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x))$ in $(\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(x)$ for $|\ell|_{\mathfrak{s}} < \gamma+2$.) This does not guarantee that the T-valued map

$$\big(\mathcal{K}^{\mathsf{M}}\boldsymbol{f}\big)(x) := \big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\boldsymbol{f}(x) + \big(\mathcal{N}^{\mathsf{M}}\boldsymbol{f}\big)(x), \quad x \in \mathbb{R}\times\mathbb{R}^d, \tag{3.15}$$

is a modelled distribution, but this turns out to be the case! Note that, unlike $\mathcal{I}$ or $\mathcal{J}^{\mathsf{M}}(x)$, the T_X-valued function $\mathcal{N}^{\mathsf{M}}\boldsymbol{f}$ is a non-local function of $\boldsymbol{f}$ – i.e., $(\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(x)$ is not a function of $\boldsymbol{f}(x)$ only. Note also that one has *formally*

$$\big(\mathcal{K}^{\mathsf{M}}\boldsymbol{f}\big)(x) = \mathcal{I}\boldsymbol{f}(x) + \sum_{|\ell|_{\mathfrak{s}}<\gamma+2} \frac{X^\ell}{\ell!}\partial^\ell\mathbf{K}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big)(x).$$

This identity gives the intuitive meaning of the polynomial part of $(\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(x)$. Decomposition (3.15) is needed to make sense of $(\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(x)$ in a rigorous way. We can prove the following theorem by the same way as [53, Theorem 5.12] except the use of a decomposition of K by integration rather than the dyadic decomposition as in [53].

Theorem 17. *Let the regularity structure $\mathcal{T}$ satisfy Assumptions* A–B1*, and let* (g,Π) *be a* $\mathbf{K}$*-admissible model on it. Let γ be a positive non-integer regularity exponent γ, and choose an integer N such that $\gamma+2<N$ and that the property* (3.3) *holds for any $|n|_{\mathfrak{s}} < N$. Then, the map $\mathcal{K}^{\mathsf{M}}$ sends continuously $\mathcal{D}^\gamma(T,\mathsf{g})$ into $\mathcal{D}^{\gamma+2}(T,\mathsf{g})$ and satisfies the intertwining identity* (3.14).

Before the proof, we recall from [53, Proposition A.1] the (anisotropic) integral Taylor formula for the remainder.

Lemma 18. *There exists a family of Borel probability measures $\{m_k\}_{k\in\mathbb{N}\times\mathbb{N}^d}$ on $[0,1]^{d+1}$ satisfying the following properties. For any smooth functions f on $\mathbb{R}\times\mathbb{R}^d$ and $\gamma>0$, one has the identity*

$$f(y) - \sum_{|k|_{\mathfrak{s}}<\gamma} \frac{\partial^k f(x)}{k!}(y-x)^k = \sum_{|\ell|_{\mathfrak{s}}\geq\gamma} \frac{(y-x)^\ell}{\ell!}\int_{[0,1]^{d+1}} \partial^\ell f(x_t)m_\ell(dt),$$

where ℓ runs over a finite set and $x_t := (x_i + t_i(y_i - x_i))_{i=0}^d$.

Proof of Theorem 17. We use the intertwining relation (3.11) to write

$$\begin{aligned}
&(\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(y) - \widehat{\mathsf{g}_{yx}}(\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(x) \\
&\quad= (\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(y) - \widehat{\mathsf{g}_{yx}}\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\boldsymbol{f}(x) - \widehat{\mathsf{g}_{yx}}(\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(x) \\
&\quad= (\mathcal{K}^{\mathsf{M}}\boldsymbol{f})(y) - \big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(y)\big)\widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x) - \widehat{\mathsf{g}_{yx}}(\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(x) \\
&\quad= \mathcal{I}\big(\boldsymbol{f}(y) - \widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big) + \mathcal{J}^{\mathsf{M}}(y)\big(\boldsymbol{f}(y) - \widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big) \\
&\qquad + \big((\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(y) - \widehat{\mathsf{g}_{yx}}(\mathcal{N}^{\mathsf{M}}\boldsymbol{f})(x)\big).
\end{aligned}$$

For the $\mathcal{I}$ term, from the continuity of $\mathcal{I}$, one has the estimate

$$\big\|\mathcal{I}\big(\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big)\big\|_\beta \lesssim \big\|\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big\|_{\beta-2} \leq \|\boldsymbol{f}\|_{\mathcal{D}^\gamma}\, d(y,x)^{\gamma+2-\beta}$$

for any $\beta \in A$. The $\mathcal{J}^{\mathsf{M}}$ and $\mathcal{N}^{\mathsf{M}}$ terms take values in the polynomial part T_X of T. Decompose $K(x,y)$ into the integral of $\tilde{q}_r$ by (3.4), and let

$$\mathcal{J}^{\mathsf{M}} =: \int_0^1 \mathcal{J}_r^{\mathsf{M}} dr \quad \text{and} \quad \mathcal{N}^{\mathsf{M}} =: \int_0^1 \mathcal{N}_r^{\mathsf{M}} dr$$

stand for the corresponding operators. Since q_1 has a smoothing property, we replace $\tilde{q}_r$ with q_r in the following calculations. Fix $n \in \mathbb{N}\times\mathbb{N}^d$, and write $(\tau)_{X^n}$ for the component of $\tau \in T$ in the direction of X^n. We have for

$$(\ominus)_r := \Big(\mathcal{J}_r^{\mathsf{M}}(y)\big(\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big) + \big(\mathcal{N}_r^{\mathsf{M}}\boldsymbol{f}\big)(y) - \widehat{\mathrm{g}_{yx}}(\mathcal{N}_r^{\mathsf{M}}\boldsymbol{f})(x)\Big)_{X^n}$$

the two decompositions

$$\begin{aligned}
(\ominus)_r = &\sum_{\beta\in A, |n|_{\mathfrak{s}}<\beta+2} \frac{1}{n!}\big\langle \Pi_y^{\mathrm{g}}\big(\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big)_\beta, \partial_y^n q_r(y,\cdot)\big\rangle \\
&+ \Big\{\frac{1}{n!}\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_y^{\mathrm{g}}\boldsymbol{f}(y), \partial_y^n q_r(y,\cdot)\big\rangle \\
&\quad - \sum_{|k|_{\mathfrak{s}}<\gamma+2-|n|_{\mathfrak{s}}} \frac{(y-x)^k}{k!}\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x), \partial_x^{n+k} q_r(x,\cdot)\big\rangle\Big\} \\
&=: (*)_r^1 + (*)_r^2
\end{aligned}$$

and

$$\begin{aligned}
(\ominus)_r &= \sum_{\beta\in A, |n|_{\mathfrak{s}}<\beta+2} \frac{1}{n!}\big\langle \Pi_y^{\mathrm{g}}\big(\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big)_\beta, \partial_y^n q_r(y,\cdot)\big\rangle \\
&\quad + \frac{1}{n!}\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x), (\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}\big\rangle + \frac{1}{n!}\big\langle \Pi_x^{\mathrm{g}}\boldsymbol{f}(x) - \Pi_y^{\mathrm{g}}\boldsymbol{f}(y), \partial_y^n q_r(y,\cdot)\big\rangle \\
&= \frac{1}{n!}\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \Pi_x^{\mathrm{g}}\boldsymbol{f}(x), (\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}\big\rangle \\
&\quad - \sum_{\beta\in A, |n|_{\mathfrak{s}}\geq\beta+2} \frac{1}{n!}\big\langle \Pi_y^{\mathrm{g}}\big(\boldsymbol{f}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big)_\beta, \partial_y^n q_r(y,\cdot)\big\rangle \\
&=: (\bigstar)_r^1 + (\bigstar)_r^2,
\end{aligned}$$

where

$$(\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}(z) := \partial_y^n q_r(y,z) - \sum_{|k|_{\mathfrak{s}}<\gamma+2-|n|_{\mathfrak{s}}} \frac{(y-x)^k}{k!}\partial_x^{n+k} q_r(x,z).$$

Choose $r_0 \in (0,1]$ such that $r_0^{\frac{1}{4}} \simeq d(y,x) \wedge 1$. We use the $(\ast)$-decomposition to estimate the integral over $0 < r < r_0$, and the $(\bigstar)$-decomposition to estimate the integral over $r_0 \le r \le 1$.

• For $r \in (0, r_0]$, we have from the bound (3.2) the estimate

$$\int_0^{r_0} \big|(\ast)_r^1\big| dr \lesssim \sum_{\beta\in A, |n|_{\mathfrak{s}}<\beta+2} d(y,x)^{\gamma-\beta} \int_0^{r_0} r^{\frac{\beta-|n|_{\mathfrak{s}}-2}{4}} dr$$
$$\lesssim \sum_{\beta\in A, |n|_{\mathfrak{s}}<\beta+2} d(y,x)^{\gamma-\beta} r_0^{\frac{\beta-|n|_{\mathfrak{s}}+2}{4}} \lesssim d(y,x)^{\gamma+2-|n|_{\mathfrak{s}}}.$$

Since $|n|_{\mathfrak{s}} < \gamma+2$, from the bound (2.31) in the reconstruction theorem, we get

$$\int_0^{r_0} \big|(\ast)_r^2\big| dr \lesssim \int_0^{r_0} r^{\frac{\gamma-|n|_{\mathfrak{s}}-2}{4}} dr + \sum_{|k|_{\mathfrak{s}}<\gamma+2-|n|_{\mathfrak{s}}} d(y,x)^{|k|_{\mathfrak{s}}} \int_0^{r_0} r^{\frac{\gamma-|n|_{\mathfrak{s}}-|k|_{\mathfrak{s}}-2}{4}} dr$$
$$\lesssim r_0^{\frac{\gamma-|n|_{\mathfrak{s}}+2}{4}} + \sum_{|k|_{\mathfrak{s}}<\gamma+2-|n|_{\mathfrak{s}}} d(y,x)^{|k|_{\mathfrak{s}}} r_0^{\frac{\gamma-|n|_{\mathfrak{s}}-|k|_{\mathfrak{s}}+2}{4}} \lesssim d(y,x)^{\gamma+2-|n|_{\mathfrak{s}}}.$$

• To deal with the integral over $r \in (r_0, 1]$, we use the $(\bigstar)$-decomposition. Since this integral does not make sense if $r_0 \ge 1$, we assume $d(y,x) \le 1$. For $(\bigstar)_r^1$, we apply Lemma 18 to write

$$(\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}(z) = \sum_{\gamma+2-|n|_{\mathfrak{s}}<|\ell|_{\mathfrak{s}}} \frac{(y-x)^\ell}{\ell!} \int_{[0,1]^{d+1}} \partial^{n+\ell} q_r(x_t, z) m_\ell(dt), \qquad (3.16)$$

where ℓ *runs over a finite set*. Note that no index n with $\gamma+2-|n|_{\mathfrak{s}} = |\ell|_{\mathfrak{s}}$ exists because $\gamma \notin \mathbb{Z}$. By decomposing

$$\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x) = \mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_{x_t}^{\mathrm{g}} \boldsymbol{f}(x_t) + \Pi_{x_t}^{\mathrm{g}} \big(\boldsymbol{f}(x_t) - \widehat{\mathrm{g}_{x_t x}} \boldsymbol{f}(x)\big),$$

and using the bounds (2.31) and (3.2), we have

$$\int_{r_0}^1 \big|(\bigstar)_r^1\big| dr$$
$$\lesssim \sum_{\gamma+2-|n|_{\mathfrak{s}}<|\ell|_{\mathfrak{s}}} d(y,x)^{|\ell|_{\mathfrak{s}}} \int_{r_0}^1 \Big\{ r^{\frac{\gamma-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}-2}{4}} + \sum_{\beta\in A, \beta<\gamma} d(y,x)^{\gamma-\beta} r^{\frac{\beta-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}-2}{4}} \Big\} dr$$
$$\lesssim \sum_{\gamma+2-|n|_{\mathfrak{s}}<|\ell|_{\mathfrak{s}}} d(y,x)^{|\ell|_{\mathfrak{s}}} \Big\{ r_0^{\frac{\gamma-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}+2}{4}} + \sum_{\beta\in A, \beta<\gamma} d(y,x)^{\gamma-\beta} r_0^{\frac{\beta-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}+2}{4}} \Big\}$$
$$\lesssim d(y,x)^{\gamma+2-|n|_{\mathfrak{s}}}.$$

We obtain the same bound for the $(\bigstar)_r^2$-term by a similar argument. Note that the terms with integer β can be excluded. Indeed, the only elements of $\mathcal{B}$ with integer homogeneity are the polynomials, and $\langle \Pi_y^{\mathrm{g}} X^k, \partial_y^n q_r(y,\cdot)\rangle = 0$ for any $|k|_{\mathfrak{s}} < N$ by the property (3.3). Therefore,

$$\int_{r_0}^1 \big|(\bigstar)_r^2\big| dr \lesssim \sum_{\beta\in A,\,|n|_{\mathfrak{s}}>\beta+2} d(y,x)^{\gamma-\beta} \int_{r_0}^1 r^{\frac{\beta-|n|_{\mathfrak{s}}-2}{4}} dr$$
$$\lesssim \sum_{\beta\in A,\,|n|_{\mathfrak{s}}>\beta+2} d(y,x)^{\gamma-\beta} r_0^{\frac{\beta-|n|_{\mathfrak{s}}+2}{4}} \lesssim d(y,x)^{\gamma+2-|n|_{\mathfrak{s}}}.$$

To show the intertwining identity (3.14), it is sufficient to obtain the estimate

$$\left|\left\langle \mathbf{K}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big) - \sum_{|n|_{\mathfrak{s}}<\gamma+2} \frac{(\cdot - x)^n}{n!} (\partial^n \mathbf{K})\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x), p_t(x,\cdot)\right\rangle\right| \lesssim t^{\frac{\gamma+2}{4}} \tag{3.17}$$

for any $\boldsymbol{f} \in \mathcal{D}^\gamma(F,\mathrm{g})$. Then, the uniqueness of the reconstruction operator gives

$$\mathbf{K}(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}) = \mathbf{R}^{\mathsf{M}}(\mathcal{K}^{\mathsf{M}} \boldsymbol{f}).$$

To prove (3.17), we write

$$\begin{aligned}
&\mathbf{K}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big) - \Pi_x^{\mathrm{g}}(\mathcal{K}^{\mathsf{M}} \boldsymbol{f})(x)\\
&\quad= \mathbf{K}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big) - \Pi_x^{\mathrm{g}}\big(\mathcal{I} + \mathcal{J}^{\mathsf{M}}(x)\big)\boldsymbol{f}(x) - \Pi_x^{\mathrm{g}}(\mathcal{N}^{\mathsf{M}} \boldsymbol{f})(x)\\
&\quad= \mathbf{K}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big) - \mathbf{K}\big(\Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big) - \Pi_x^{\mathrm{g}}(\mathcal{N}^{\mathsf{M}} \boldsymbol{f})(x)\\
&\quad= \mathbf{K}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big) - \sum_{|n|_{\mathfrak{s}}<\gamma+2} \frac{(\cdot - x)^n}{n!} (\partial^n \mathbf{K})\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x).
\end{aligned}$$

We decompose $\mathbf{K}$ into the integral of K_r over $r \in [0,1]$ by (3.4) and ignore $P(\partial)K_1$ as it has a smoothing property. Since $\int_{\mathbb{R}\times\mathbb{R}^d} q_r(y,\cdot)p_t(x,y)dy = q_{r+t}(x,\cdot)$ by definition, we have

$$\begin{aligned}
&\left|\left\langle K_r\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big) - \sum_{|n|_{\mathfrak{s}}<\gamma+2} \frac{(\cdot - x)^n}{n!} (\partial^n K_r)\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x), p_t(x,\cdot)\right\rangle\right|\\
&\quad\lesssim \big|K_{r+t}\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x)\big| + \sum_{|n|_{\mathfrak{s}}<\gamma+2} t^{\frac{|n|_{\mathfrak{s}}}{4}} \big|(\partial^n K_r)\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)\big)(x)\big|\\
&\quad\lesssim (r+t)^{\frac{\gamma-2}{4}} + \sum_{|n|_{\mathfrak{s}}<\gamma+2} t^{\frac{|n|_{\mathfrak{s}}}{4}} r^{\frac{\gamma-|n|_{\mathfrak{s}}-2}{4}}.
\end{aligned}$$

Since $\frac{\gamma-|n|_{\mathfrak{s}}-2}{4} > -1$, the integral over $r \in [0,t]$ gives the upper bound $t^{\frac{\gamma+2}{4}}$. For the integral over $r \in [t,1]$, we use the representation $(\bigstar)_r^1$ with

$$n = 0.$$

By the estimate of $(\bigstar)_r^1$ obtained before, we have

$$\begin{aligned}
&\left|\left\langle (\bigstar)_r^1(\cdot), p_t(x,\cdot)\right\rangle\right| \\
&\lesssim \sum_{\gamma+2<|\ell|_{\mathfrak{s}}} \int_{\mathbb{R}\times\mathbb{R}^d} d(y,x)^{|\ell|_{\mathfrak{s}}} \left\{ r^{\frac{\gamma-|\ell|_{\mathfrak{s}}-2}{4}} + \sum_{\beta\in A,\,\beta<\gamma} d(y,x)^{\gamma-\beta} r^{\frac{\beta-|\ell|_{\mathfrak{s}}-2}{4}} \right\} p_t(x,y)dy \\
&\lesssim \sum_{\gamma+2<|\ell|_{\mathfrak{s}}} \left\{ t^{\frac{|\ell|_{\mathfrak{s}}}{4}} r^{\frac{\gamma-|\ell|_{\mathfrak{s}}-2}{4}} + \sum_{\beta\in A,\,\beta<\gamma} t^{\frac{\gamma-\beta+|\ell|_{\mathfrak{s}}}{4}} r^{\frac{\beta-|\ell|_{\mathfrak{s}}-2}{4}} \right\}.
\end{aligned}$$

Then, the integral over $r \in [t, 1]$ gives the upper bound $t^{\frac{\gamma+2}{4}}$. ■

Note that the intertwining relation (3.14) between $\mathbf{K}$ and $\mathcal{K}^{\mathsf{M}}$ provides indeed an "upgraded" version of the defining identity (3.10) for a $\mathbf{K}$-admissible model in so far as the former reduces to the latter when applied to the modelled distribution

$$\boldsymbol{f}(x) = \boldsymbol{h}^\tau(x) = \sum_{\sigma<\tau} \mathrm{g}_x(\tau/\sigma)\sigma.$$

Indeed, on the one hand, we have $\mathbf{R}^{\mathsf{M}}\boldsymbol{h}^\tau = \Pi\tau$. On the other hand, $\mathcal{K}^{\mathsf{M}}\boldsymbol{h}^\tau$ has positive regularity and takes its values in a function-like sector when the model takes values in the space of continuous functions, so Corollary 6 applies and identifies the reconstruction of $\mathcal{K}^{\mathsf{M}}\boldsymbol{h}^\tau$ as $\Pi_x^{\mathrm{g}}(\mathcal{K}^{\mathsf{M}}\boldsymbol{h}^\tau(x))(x)$, equal to $\Pi(\mathcal{I}\tau)$, as all the x-indexed polynomial terms are null when evaluated at x.

3.5. Building admissible smooth models

We left aside in Section 3.3 the non-elementary question of existence of non-trivial admissible models to concentrate on their properties. We construct in this section a large class of admissible models for which all $(\Pi\tau)_{\tau\in T}$ and $(\mathrm{g}(\sigma))_{\sigma\in T^+}$ are smooth functions. In applications to singular stochastic PDEs, such models can be built from realizations of the noise(s) in the equation.

Recall that Assumption B1 in Section 3.2 describes the action of the "recentering operator" Δ on elements of T of the form $\mathcal{I}_k\tau$. We single out for our needs an assumption on Δ that provides a crucial induction structure.

Assumption B2. There exists an increasing sequence $\{\mathcal{B}^{(n)}\}_{n=0}^\infty$ of subsets of $\mathcal{B}$ such that $\mathcal{B}^{(0)} = \{X^k\}_{k\in\mathbb{N}\times\mathbb{N}^d}$, $\mathcal{B} = \bigcup_{n=0}^\infty \mathcal{B}^{(n)}$, and

$$\Delta\tau - \tau\otimes\mathbf{1} \in T^{(n-1)} \otimes T^{(n-1)+}$$

for any n and $\tau\in\mathcal{B}^{(n)}$, where $T^{(n-1)}$ is the vector space spanned by $\mathcal{B}^{(n-1)}$, and $T^{(n-1)+}$ is the subalgebra of T^+ generated by the symbols

$$\{X_+^{e_i}\}_{0\le i\le d} \cup \left\{\mathcal{I}_k^+\sigma\right\}_{\sigma\in\mathcal{B}^{(n-1)},\,k\in\mathbb{N}\times\mathbb{N}^d}.$$

Here, $\mathcal{B}^{(-1)} := \emptyset$.

Assumptions B1 and B2 are jointly called Assumption B. Assumption B2 is satisfied by the regularity structures of decorated trees as defined in Section 9. Indeed, for this case, we can define $\mathcal{B}^{(n)}$ as the set of decorated trees τ such that the sum of the number of edges in τ and the total number of $\mathfrak{n}$ decorations is n. That is, n means the "complicatedness" of the tree. It should be noted that, by the definition (3.7), we have

$$\Delta^+ \mathcal{I}_k^+ \tau \in T^{(n)+} \otimes T^{(n)+}$$

for any $\tau \in \mathcal{B}^{(n)}$, and thus, Δ^+ is closed within $T^{(n)+}$. Therefore, the pair of subspaces

$$\mathfrak{T}^{(n)} := \big(T^{(n)+}, T^{(n)}\big),$$

equipped with the restrictions of the Δ^+ and Δ maps, is also a regularity structure.

Under Assumptions A–B, formula (3.13) in the proof of Proposition 15 shows that the g-map of an admissible model (Π, g) is uniquely determined by its Π-map. The following theorem essentially comes from [53, Proposition 3.32], which is a generalization of the "Lyons' extension theorem".

Theorem 19. *Let* $\mathfrak{T}$ *be a regularity structure satisfying Assumption* A–B *and* (3.1)*. One can associate to any family* $([\tau]; \tau \in \mathcal{B}, |\tau| < 0)$ *of smooth functions on* $\mathbb{R} \times \mathbb{R}^d$ *a unique* $\mathbf{K}$*-admissible model* (g, Π) *on* $\mathfrak{T}$ *such that* $\Pi\tau = [\tau]$ *for all* $\tau \in \mathcal{B}$ *with* $|\tau| < 0$.

Proof. We set the scene for an inductive proof of the statement, taking advantage of the induction structure given by Assumption B2. We will define inductively on $n \in \mathbb{N}$ the $\mathbf{K}$-admissible models $\mathsf{M}^{(n)} = (\mathrm{g}^{(n)}, \Pi^{(n)})$ on $\mathfrak{T}^{(n)}$ over such that

$$\mathrm{g}^{(n)} : T^{(n)+} \to \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^d), \quad \Pi^{(n)} : T^{(n)} \to \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^d)$$

and

$$\mathrm{g}^{(m)}|_{T^{(n)+}} = \mathrm{g}^{(n)}, \quad \Pi^{(m)}|_{T^{(n)}} = \Pi^{(n)}$$

for any $m > n$. We denote by $\mathbf{R}^{\mathsf{M}^{(n)}}$ the reconstruction operator associated with the model $\mathsf{M}^{(n)}$. Note that the model $\mathsf{M}^{(0)}$ on $\mathfrak{T}^{(0)}$ is canonically defined by Assumption A2.

• We now define an extension $\mathsf{M}^{(n)}$ of $\mathsf{M}^{(n-1)}$ on $\mathfrak{T}^{(n)}$. It is sufficient to define $\Pi^{(n)}\tau$ and $\mathrm{g}^{(n)}(\mathcal{I}_k^+\tau)$ for $\tau \in \mathcal{B}^{(n)}$. Recall from the sentences after Theorem 4 that the function

$$\boldsymbol{h}^\tau(x) := \sum_{\sigma < \tau} \mathrm{g}_x^{(n-1)}(\tau/\sigma)\, \sigma$$

is an element of $\mathcal{D}^{|\tau|}(T^{(n-1)}, \mathrm{g}^{(n-1)})$ and its reconstruction is a candidate of $\Pi^{(n)}\tau$. Given that $\mathrm{g}^{(n-1)}$ and $\Pi^{(n-1)}$ take values in smooth functions, any smooth function is a reconstruction of $\boldsymbol{h}^\tau$ for the model $\mathsf{M}^{(n-1)}$, if $|\tau| < 0$. (Recall that the reconstruction operator is defined uniquely only when acting on modelled distributions of positive regularity. We are working here with a modelled distribution of negative regularity when $|\tau| < 0$.) Define

$$\Pi^{(n)}\tau = \begin{cases} [\tau] & (|\tau| < 0), \\ \mathbf{R}^{\mathsf{M}^{(n-1)}}(\boldsymbol{h}^\tau) & (|\tau| > 0). \end{cases}$$

This is a smooth function in both cases. (Recall that $\mathbf{1}$ is the only element of T of null homogeneity by Assumption A1.) The map $\Pi^{(n)}$ coincides with $\Pi^{(n-1)}$ on $T^{(n-1)}$.

• Define then an extension $\mathsf{g}^{(n)}$ of $\mathsf{g}^{(n-1)}$ to $T^{(n)+}$ by requiring that it is multiplicative, and by setting

$$\mathsf{g}_x^{(n)}(\mathcal{I}_k^+\tau) := \sum_{\sigma\le\tau;|k|_{\mathfrak{s}}<|\sigma|+2} \mathsf{g}_x^{(n-1)}(\tau/\sigma)\,\partial^k\mathbf{K}\Big(\big(\Pi^{(n)}\big)_x^{\mathsf{g}^{(n-1)}}\sigma\Big)(x),$$

for all $\tau\in\mathcal{B}^{(n)}$, in view of (3.13). Note that $(T^{(n-1)+},T^{(n)})$ is a regularity structure and $(\mathsf{g}^{(n-1)},\Pi^{(n)})$ is a model over it. Closing the induction step amounts to proving that

$$\big|\mathsf{g}_{yx}^{(n)}(\mathcal{I}_k^+\tau)\big| \lesssim d(y,x)^{|\tau|+2-|k|_{\mathfrak{s}}} \tag{3.18}$$

for every $k\in\mathbb{N}\times\mathbb{N}^d$ with $|k|_{\mathfrak{s}}<|\tau|+2$. Look for that purpose at the $T^{(n-1)+}$-valued function

$$\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau := \big(\mathcal{I}^+ + \mathcal{J}^{\mathsf{M}^{(n-1)}+}(x)\big)\boldsymbol{h}^\tau(x) + (\mathcal{N}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau)(x),$$

where $\mathcal{I}^+:\mathcal{B}\to\mathcal{B}^+$ is linearly extended by imposing $\mathcal{I}^+(X^k)=0$, the linear map $\mathcal{J}^{\mathsf{M}^{(n-1)}+}:T^{(n-1)}\to T^+$ is defined by

$$\mathcal{J}^{\mathsf{M}^{(n-1)}+}\sigma := \sum_{|k|_{\mathfrak{s}}<|\tau|+2}\frac{X_+^k}{k!}\partial^k\mathbf{K}\Big(\big(\Pi^{(n-1)}\big)_x^{\mathsf{g}^{(n-1)}}\sigma\Big)(x),\qquad \sigma\in\mathcal{B}^{(n-1)}$$

in the same way as $\mathcal{J}^{\mathsf{M}}$ by replacing M with $\mathsf{M}^{(n-1)}$ and X^k with X_+^k, and the function $\mathcal{N}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau$ is defined by

$$(\mathcal{N}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau)(x) := \sum_{|k|_{\mathfrak{s}}<|\tau|+2}\frac{X_+^k}{k!}\partial^k\mathbf{K}\Big(\mathbf{R}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau(x) - \big(\Pi^{(n-1)}\big)_x^{\mathsf{g}^{(n-1)}}\boldsymbol{h}^\tau(x)\Big)(x)$$

in the same way as $\mathcal{N}^{\mathsf{M}}$ by replacing M with $\mathsf{M}^{(n-1)}$ and X^k with X_+^k, where $\mathbf{R}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau$ is defined as $\Pi^{(n)}\tau$. Then, in the same proof as Theorem 17, we can prove that $\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau$ is an element of $\mathcal{D}^{|\tau|+2}(T^{(n-1)+},\mathsf{g}^{(n-1)})$ and that, denoting by $\mathbf{R}^{\mathsf{g}^{(n-1)}}$ the reconstruction operator associated with the model $(\mathsf{g}^{(n-1)},\mathsf{g}^{(n-1)})$ on $(T^{(n-1)+},T^{(n-1)+})$, we have $\mathbf{R}^{\mathsf{g}^{(n-1)}}\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau = \mathbf{K}(\mathbf{R}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau)$. In the proof, Proposition 16 has the role of Corollary 14. Since

$$\mathbf{R}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau(x) - (\Pi^{(n-1)})_x^{\mathsf{g}^{(n-1)}}\boldsymbol{h}^\tau(x) = (\Pi^{(n)})_x^{\mathsf{g}^{(n-1)}}\tau$$

in the definition of $(\mathcal{N}^{\mathsf{M}^{(n-1)}}\boldsymbol{h}^\tau)(x)$, we have

$$\big(\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau\big)(x) = \mathcal{I}^+\big(\boldsymbol{h}^\tau(x)\big) + \sum_{|k|_{\mathfrak{s}}<|\tau|+2}\mathsf{g}_x^{(n)}(\mathcal{I}_k^+\tau)\frac{X_+^k}{k!}.$$

The X_+^k-component of

$$\big(\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau\big)(y) - \widehat{\mathsf{g}_{yx}^{(n-1)}}{}^+\big(\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau\big)(x)$$

is then equal to

$$\mathsf{g}_y^{(n)}(\mathcal{I}_k^+\tau) - \sum_{\eta<\tau}\mathsf{g}_x^{(n)}(\tau/\eta)\mathsf{g}_{yx}^{(n)}(\mathcal{I}_k^+\eta) - \sum_m \mathsf{g}_x^{(n)}(\mathcal{I}_{k+m}^+\tau)\frac{(y-x)^m}{m!} = \mathsf{g}_{yx}^{(n)}(\mathcal{I}_k^+\tau),$$

and of size $d(y,x)^{|\tau|+2-|k|_{\mathfrak{s}}}$, since $\mathcal{K}^{\mathsf{M}^{(n-1)}+}\boldsymbol{h}^\tau \in \mathcal{D}^{|\tau|+2}(T^{(n-1)+}, \mathsf{g}^{(n-1)})$. This shows the bound (3.18).

• It remains to show that Π is $\mathbf{K}$-admissible. Given that we assume $\beta_0 = \min A > -2$, the elements of T of the form $\mathcal{I}\tau$ have positive homogeneity. So, the definition of Π on $\mathcal{I}\tau$ comes under the form of the reconstruction of a modelled distribution $\boldsymbol{h}^{\mathcal{I}\tau}$. Since $\boldsymbol{h}^{\mathcal{I}\tau}$ is function-like with the $\mathbf{1}$-component $\mathsf{g}_x(\mathcal{I}^+\tau)$, by Corollary 7, it follows that

$$\Pi(\mathcal{I}\tau)(x) = \mathbf{R}^{\mathsf{M}}(\boldsymbol{h}^{\mathcal{I}\tau})(x) = \mathsf{g}_x(\mathcal{I}^+\tau) = \mathbf{K}(\Pi\tau)(x).$$

• The above construction makes it clear that the map Π is entirely determined from its restriction to the elements of negative homogeneity. The uniqueness part of the statement of the theorem follows then from formula (3.13) giving $\mathsf{g}_x(\mathcal{I}_k^+\tau)$ as it shows that the map g is entirely determined by the Π map under the Assumption B1–B2. ■

4. Solving singular PDEs within regularity structures

In this section, we formulate singular stochastic PDEs in the sense of modelled distributions. We trade in this section the generality of the above results for the simplicity of an example that contains the main difficulties of the general case. The reader can consult [53] or [17] for a description of the general case. We consider the generalized (KPZ) equation

$$\begin{aligned}\big(\partial_{x_0} - \Delta_{x'} + 1\big)u &= f(u)\zeta + \sum_{i,j=1}^d g_2^{ij}(u)(\partial_{x_i}u)(\partial_{x_j}u)\\ &\quad + \sum_{i=1}^d g_1^i(u)(\partial_{x_i}u) + g_0(u)\\ &=: f(u)\zeta + g_2(u)(\partial_{x'}u)^2 + g_1(u)\partial_{x'}u + g_0(u)\\ &=: f(u)\zeta + g(u,\partial_{x'}u) \qquad (4.1)\end{aligned}$$

with a noise $\zeta \in \mathcal{C}^{\beta_0}$. (Remember that the minimum homogeneity in a regularity structure associated with a singular stochastic PDE coincides with the minimum of the regularities of the noises in the equation.) This type of equation appears in a number of problems.

If $d = 1$ and ζ is a space-time white noise, then (4.1) contains the KPZ equation, which appears in the large-scale picture of one-dimensional random interface evolutions. Here, u is scalar valued but a vector-valued case is used in the description of the random motion of a rubber on a manifold [55], a random perturbation of the harmonic flow map on loops. If $d = 2, 3$ and ζ is a space white noise, then (4.1) contains the generalized PAM

$$(\partial_t - \Delta_x)u = f(u)\zeta.$$

The differential equation (4.1) with the initial value $u(0,x') = u_0(x')$ has an equivalent integral form

$$u(x) = e^{x_0(\Delta_{x'}-1)}u_0(x') + \mathbf{L}^{-1}\big(f(u)\zeta + g(u, \partial_{x'}u)\big)(x).$$

Under an appropriate setting, the generalized (KPZ) equation (4.1) will be lifted to the following equation on modelled distributions $\boldsymbol{v} \in \mathcal{D}^\gamma(T, \mathrm{g})$:

$$\boldsymbol{v} = \boldsymbol{h} + \mathcal{P}^{\mathsf{M}}\big(f^\star(\boldsymbol{v})\Xi + g^\star(\boldsymbol{v}, D\boldsymbol{v})\big) \tag{4.2}$$

for some T_X-valued modelled distribution $\boldsymbol{h}$ and an operator $\mathcal{P}^{\mathsf{M}}$ having the role of $\mathbf{L}^{-1}$. This section is dedicated to giving the meaning to the equation (4.2) and showing that this equation has a unique solution on a small time interval $(0, t_0)$; this is the content of Theorem 23, which is the main result of this section. At the end of this section, we will be in a position to define the model-dependent solution of the singular equation (4.1) as the model-dependent reconstruction $u^{\mathsf{M}} = \mathsf{R}^{\mathsf{M}}(\boldsymbol{u}^{\mathsf{M}})$ of the unique solution $\boldsymbol{u}^{\mathsf{M}}$ of (4.2). The function u^{M} will then appear as a continuous function of M.

The restriction to each band $[0, t_0] \times [-R, R]$ of spacetime white noise has a norm growing indefinitely as R goes to infinity for each fixed $t_0 > 0$. To avoid working with unbounded spatial domains and functional spaces involving spatial weights, we will assume that all the objects are $\mathbb{Z}$-periodic in space – they would be $\mathbb{Z}^d$-periodic in space in a more general setting. The function $\boldsymbol{h}$ in (4.2) plays the role of the regularity structure lift of the propagator of the initial condition u_0. The use of time weights to take care of the free propagation $(e^{t(\Delta-1)}u_0)_{t>0}$ of the initial condition in a regularity structures setting is made necessary by the classical sharp estimate

$$\big\|\partial_x^k e^{t(\Delta-1)}u_0\big\|_\infty \lesssim t^{-\frac{(|k|_{\mathfrak{s}}-\alpha)\vee 0}{2}} \|u_0\|_{\mathcal{C}^\alpha(\mathbb{R}^d)}. \tag{4.3}$$

Theorem 23 is proved under spatial periodic boundary conditions and in the space of modelled distributions involving temporal weights exploding in $t = 0^+$. We introduce the former in Section 4.1 and the latter in Section 4.2. We examine in Section 4.3 the notion of non-anticipative operator, involved in the analysis of equation (4.2). We prove in Section 4.4 that (4.2) is locally in time well-posed.

4.1. Spatially periodic models

We work on the models and modelled distributions that are spatially $\mathbb{Z}^d$-periodic, with $d = 1$ here – we give the definitions for an arbitrary space dimension d. All the results and estimates proved above hold true in the periodic case. For any $x = (x_0, x') \in \mathbb{R} \times \mathbb{R}^d$ and $m \in \mathbb{Z}^d$, denote by

$$x + m := (x_0, x' + m).$$

Definition. A *model* $\mathsf{M} = (g, \Pi)$ is said to be $\mathbb{Z}^d$*-periodic* if for any $m \in \mathbb{Z}^d$,

$$\mathrm{g}_{y+m,x+m} = \mathrm{g}_{yx}, \quad \langle \Pi^{\mathrm{g}}_{x+m}\tau, \varphi(\cdot + m)\rangle = \langle \Pi^{\mathrm{g}}_x \tau, \varphi(\cdot)\rangle,$$

for all $x, y \in \mathbb{R} \times \mathbb{R}^d$, $\tau \in T$ and all $\varphi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$.

The canonical model (g, Π) on the polynomial regularity structure (T_X^+, T_X) is $\mathbb{Z}^d$-periodic in the above sense. Note that $\mathrm{g}_x(X_+^n) = x^n$ and $(\Pi X^n)(x) = x^n$ are *not* $\mathbb{Z}^d$-periodic functions. This is the reason why we do not impose periodic conditions on g and Π. It is elementary to see that if M is a $\mathbb{Z}^d$-periodic model on $\mathcal{T}$ and $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$ is $\mathbb{Z}^d$-periodic, with γ positive, then $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}$ is also $\mathbb{Z}^d$-periodic, in the sense that

$$\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, \varphi(\cdot + m)\rangle = \langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, \varphi(\cdot)\rangle,$$

for all $m \in \mathbb{Z}^d$ and $\varphi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$ – see [53, Proposition 3.38]. *All objects in the remainder of this section are implicitly assumed to be* $\mathbb{Z}^d$*-periodic.*

4.2. Modelled distributions with singularity at $x_0 = 0$

We use time weights

$$\omega(x) := |x_0|^{\frac{1}{2}} \wedge 1, \quad \omega(x, y) := \omega(x) \wedge \omega(y)$$

to treat the boundary condition at $x_0 = 0$.

Definition. Fix two exponents $\eta \leq \gamma \in \mathbb{R}$. One defines the space $\mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ of *modelled distributions with singularity of weight* η *at* $x_0 = 0$ as the space of functions $\boldsymbol{f}$ from $\mathbb{R} \times \mathbb{R}^d \setminus \{x_0 = 0\}$ into $T_{<\gamma}$ such that

$$\begin{aligned} [\![\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}} &:= \max_{\beta<\gamma} \sup_{x \in (\mathbb{R}\setminus\{0\})\times\mathbb{R}^d} \frac{\|\boldsymbol{f}(x)\|_\beta}{\omega(x)^{(\eta-\beta)\wedge 0}} < \infty, \\ \|\boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}} &:= \max_{\beta<\gamma} \sup_{x,y \in (\mathbb{R}\setminus\{0\})\times\mathbb{R}^d,\, d(x,y)\leq\omega(x,y)} \frac{\|\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\|_\beta}{\omega(x,y)^{\eta-\gamma} d(y,x)^{\gamma-\beta}} < \infty. \end{aligned} \tag{4.4}$$

Set $|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}^{\gamma,\eta}} := [\![\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}} + \|\boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}}$.

One also talks of singular modelled distributions. An example of singular modelled distributions is obtained as follows. Given $\eta \in \mathbb{R}$ and $v \in \mathcal{C}^\eta(\mathbb{T}^d)$, the T_X-valued function

$$(P_\gamma v)(x) := \mathbf{1}_{x_0>0} \sum_{|k|_{\mathfrak{s}}<\gamma} \partial^k \big(e^{x_0(\Delta_{x'}-1)} v\big)(x) \frac{X^k}{k!} \tag{4.5}$$

belongs to $\mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ for any $\gamma \geq \eta$. This is a consequence of the Schauder estimates satisfied by the heat semigroup recalled in (4.3) – see, e.g., [53, Lemma 7.5].

We recall some embedding theorems. It is easy to see that $\|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta'}} \leq \|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}}$ if $\eta' \leq \eta$. If $\eta \leq \gamma' \leq \gamma$, we also have

$$\|\!| \mathcal{Q}_{<\gamma'} f \|\!|_{\mathcal{D}^{\gamma',\eta}} \lesssim \|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}}$$

with an implicit constant depending on the model M – see, e.g., [65, Proposition 3.5]. Instead of $[\![f]\!]_{\mathcal{D}^{\gamma,\eta}}$ and $\|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}}$, it will be convenient to consider the seminorms

$$[\![f]\!]'_{\mathcal{D}^{\gamma,\eta}} := \max_{\beta<\gamma} \sup_{x\in(\mathbb{R}\setminus\{0\})\times\mathbb{R}^d} \frac{\| f(x)\|_\beta}{\omega(x)^{\eta-\beta}}$$

and $\|\!| f \|\!|'_{\mathcal{D}^{\gamma,\eta}} := [\![f]\!]'_{\mathcal{D}^{\gamma,\eta}} + \| f \|_{\mathcal{D}^{\gamma,\eta}}$. In general, $\|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}} \leq \|\!| f \|\!|'_{\mathcal{D}^{\gamma,\eta}}$, but the reverse inequality fails. However, for any $f \in \mathcal{D}^{\gamma,\eta}$ such that

$$\lim_{x_0\to 0} \mathcal{Q}_\beta f(x) = 0$$

for any $\beta < \eta$, the reverse inequality $\|\!| f \|\!|'_{\mathcal{D}^{\gamma,\eta}} \lesssim \|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}}$ holds with an implicit constant depending on the model M – see, e.g., [53, Lemma 6.5].

The reconstruction theorem, Theorem 4, is extended to singular modelled distributions as follows. See Appendix C.1 for a detailed proof. An extension to the inhomogeneous integral kernels can be found in [65].

Theorem 20. *Let* $\mathsf{M} = (g, \Pi)$ *be a model over* $\mathcal{T}$ *such that* $-2 < \beta_0 < 0$. *Assume that* $-2 < \eta \leq \gamma$, *with* $\gamma > 0$. *Then, there exists a continuous linear operator*

$$\mathbf{R}^{\mathsf{M}} : \mathcal{D}^{\gamma,\eta}(T, \mathrm{g}) \to \mathcal{C}^{\eta\wedge\beta_0}(\mathbb{R}\times\mathbb{R}^d)$$

such that, for any $f \in \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ *and* $n \in \mathbb{N}\times\mathbb{N}^d$, *the bound*

$$\left|\left\langle \mathbf{R}^{\mathsf{M}} f - \Pi_x^{\mathrm{g}} f(x), \partial_x^n p_t(x,\cdot)\right\rangle\right| \lesssim \|\!| f \|\!|_{\mathcal{D}^{\gamma,\eta}} \omega(x)^{(\eta\wedge\beta_0)-\gamma} t^{\frac{\gamma-|n|_{\mathfrak{s}}}{4}} \tag{4.6}$$

holds uniformly over $f \in \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$, $x \in \mathbb{R}\times\mathbb{R}^d$, *and* $0 < t \leq \omega(x)^4$, *where the implicit proportional constant depends polynomially on* $\|\mathrm{g}\|_\gamma + \|\Pi^{\mathrm{g}}\|_\gamma$ *and is independent of* f. *Such an operator is unique if the exponent* γ *is positive.*

The operators discussed in previous sections can be extended to the spaces $\mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$, as follows. All of the following maps are locally Lipschitz continuous. For the detailed proofs, see [53, Propositions 6.12, 6.13, 6.15, and 6.16]. Denote by $\mathcal{D}_\alpha^{\gamma,\eta}(T, \mathrm{g})$ the space of modelled distributions $f \in \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ of the form

$$f = \sum_{\alpha\leq|\tau|<\gamma} f_\tau \tau. \tag{4.7}$$

- (Proposition 8′) Let $\alpha_1, \alpha_2 \le 0 < \gamma_1, \gamma_2$, and set $\gamma = (\gamma_1 + \alpha_2) \wedge (\gamma_2 + \alpha_1)$ and $\eta = (\eta_1 + \alpha_2) \wedge (\eta_2 + \alpha_1) \wedge (\eta_1 + \eta_2)$. If a regular product $\star : V \times W \to T$ is given, then

$$\mathcal{D}_{\alpha_1}^{\gamma_1,\eta_1}(V,\mathrm{g}) \times \mathcal{D}_{\alpha_2}^{\gamma_2,\eta_2}(W,\mathrm{g}) \ni (\boldsymbol{f}_1, \boldsymbol{f}_2) \mapsto \mathcal{Q}_{<\gamma}(\boldsymbol{f}_1 \star \boldsymbol{f}_2) \in \mathcal{D}_{\alpha_1+\alpha_2}^{\gamma,\eta}(T,\mathrm{g}).$$

- (Proposition 9′) Let $\gamma > 0$ and $0 \le \eta \le \gamma$. If an associative regular product $\star : V \times V \to V$ and a smooth function F is given, then

$$\mathcal{D}^{\gamma,\eta}(V,\mathrm{g}) \ni \boldsymbol{f} \mapsto F^{\star}(\boldsymbol{f}) \in \mathcal{D}^{\gamma,\eta}(V,\mathrm{g}).$$

- (Proposition 10′) Let $\gamma > 1$. If a derivative $D : T \to T$ is given, then

$$\mathcal{D}^{\gamma,\eta}(T,\mathrm{g}) \ni \boldsymbol{f} \mapsto D\boldsymbol{f} \in \mathcal{D}^{\gamma-1,\eta-1}(T,\mathrm{g}).$$

- (Theorem 17′) Let $\gamma > 0$ and $-2 < \eta \wedge \beta_0$. If Π is $\mathbf{K}$-admissible,

$$\mathcal{D}^{\gamma,\eta}(T,\mathrm{g}) \ni \boldsymbol{f} \mapsto \mathcal{K}^{\mathsf{M}}\boldsymbol{f} \in \mathcal{D}^{\gamma+2,\eta\wedge\beta_0+2}(T,\mathrm{g}).$$

A sketch of the proof of Theorem 17′ is given in Appendix C.1. About this statement, note here the gain in the explosion exponent after we applied the operator $\mathcal{K}^{\mathsf{M}}$. We will use this gain in Section 4.4 to gain a small contraction factor in the fixed-point formulation of equation (4.1) as an equation on a space of modelled distributions.

4.3. Non-anticipative operators

A function f on $(\mathbb{R} \times \mathbb{T}^d)^2$ is said to be *non-anticipative* if $f((x_0, x'), (y_0, y')) = 0$, whenever $x_0 < y_0$. The kernel P of the resolution operator $\mathbf{L}^{-1}$ (in the sense that $\mathbf{L}^{-1} f(x) = \int_{\mathbb{R}\times\mathbb{T}^d} P(x-y) f(y)dy$) is of the form

$$P(x,y) = \mathbf{1}_{x_0>y_0}\, p_{x_0-y_0}(x'-y'),$$

where p_t is the kernel of $e^{t(\Delta-1)}$; thus, P is non-anticipative. The aim of this section is to prove the refined multilevel Schauder estimate (Proposition 22) associated with the non-anticipative operator $\mathbf{L}^{-1}$.

We consider the modelled distributions defined on the domain $(0,t) \times \mathbb{T}^d$ for a given positive time t. Denote by $\mathcal{D}_{(0,t)}^{\gamma,\eta}(T,\mathrm{g})$ the set of functions $\boldsymbol{f} : (0,t) \times \mathbb{T}^d \to T_{<\gamma}$ such that the bounds (4.4) hold with the domain of x, y restricted to $(0,t) \times \mathbb{T}^d$. Denote by

$$|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{(0,t)}^{\gamma,\eta}} := [\![\boldsymbol{f}]\!]_{\mathcal{D}_{(0,t)}^{\gamma,\eta}} + \| \boldsymbol{f} \|_{\mathcal{D}_{(0,t)}^{\gamma,\eta}}$$

the associated norms. It is also useful to consider $[\![\boldsymbol{f}]\!]'_{\mathcal{D}_{(0,t)}^{\gamma,\eta}}$ and $|\!|\!| \boldsymbol{f} |\!|\!|'_{\mathcal{D}_{(0,t)}^{\gamma,\eta}} := [\![\boldsymbol{f}]\!]'_{\mathcal{D}_{(0,t)}^{\gamma,\eta}} + \| \boldsymbol{f} \|_{\mathcal{D}_{(0,t)}^{\gamma,\eta}}$. Since $\omega(x), \omega(y) \le t^{\frac{1}{2}}$ if $x, y \in (0,t) \times \mathbb{T}^d$, we have

$$|\!|\!| \boldsymbol{f} |\!|\!|'_{\mathcal{D}_{(0,t)}^{\gamma,\eta-\kappa}} \lesssim t^{\frac{\kappa}{2}} |\!|\!| \boldsymbol{f} |\!|\!|'_{\mathcal{D}_{(0,t)}^{\gamma,\eta}}$$

for any $\kappa > 0$ small enough such that $[\eta - \kappa, \eta) \cap A = \emptyset$. The small factor $t^{\kappa/2}$ is used in the fixed-point problem in the next section. To apply the reconstruction operator to locally defined modelled distributions, we use the cut-off operator. The following result is obtained from Proposition 8′ and Lemma 51. See [65, Lemma 5.7] for the detailed proof.

Proposition 21. *Let* $\mathsf{M} = (g, \Pi)$ *be a model over* $\mathcal{T}$ *such that* $-2 < \beta_0 < 0$, *and let* $\gamma > 0$ *and* $\eta \le \gamma \wedge \beta_0$. *Fix a smooth non-increasing function* $\chi : (0, \infty) \to [0, 1]$ *such that* $\chi(t) = 1$ *if* $0 < t \le \frac{1}{2}$ *and* $\chi(t) = 0$ *if* $t \ge 1$. *For each* $t > 0$ *and* $x \in \mathbb{R} \times \mathbb{T}^d$, *we set* $\chi_t(x) = \mathbf{1}_{x_0 > 0}\chi(x_0/t)$ *and define*

$$\boldsymbol{\chi}_t(x) := \sum_{|k|_{\mathfrak{s}} < \gamma - \beta_0} \frac{\partial_x^k \chi_t(x)}{k!} X^k.$$

Then, one can define the linear operator $C_t : \mathcal{D}^{\gamma,\eta}_{(0,t)}(T, \mathrm{g}) \to \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ *by*

$$C_t \boldsymbol{f}(x) = \mathbf{1}_{(0,t)\times\mathbb{T}^d}(x) \mathcal{Q}_{<\gamma}\big(\boldsymbol{\chi}_t(x) \star \boldsymbol{f}(x)\big),$$

and C_t *is uniformly bounded over* $t \in (0, 1]$ *and satisfies* $(C_t \boldsymbol{f})|_{(0,\frac{t}{2})\times\mathbb{T}^d} = \boldsymbol{f}|_{(0,\frac{t}{2})\times\mathbb{T}^d}$.

Proposition 22. *Pick* $\gamma > 0$, $-2 < \eta \le \beta_0 < 0$, *and* $0 < \rho \le \gamma + 2$. *For any* $\mathbf{K}$*-admissible model* $\mathsf{M} = (g, \Pi)$ *and* $t \in (0, 1]$, *there exists a continuous linear map* $\mathcal{P}^{\mathsf{M}}_t : \mathcal{D}^{\gamma,\eta}_{(0,t)}(T, \mathrm{g}) \to \mathcal{D}^{\rho,\eta+2}(T, \mathrm{g})$ *such that the following properties hold for any* $\boldsymbol{f} \in \mathcal{D}^{\gamma,\eta}_{(0,t)}(T, \mathrm{g})$.

(1) *For any* $x \in (0, \frac{t}{2}) \times \mathbb{T}^d$, *one has* $\mathcal{P}^{\mathsf{M}}_t \boldsymbol{f}(x) - \mathcal{I}\boldsymbol{f}(x) \in T_X$.

(2) *For any* $\kappa > 0$, *one has*

$$|\!|\!| \mathcal{P}^{\mathsf{M}}_t \boldsymbol{f} |\!|\!|_{\mathcal{D}^{\rho,\eta+2-\kappa}_{(0,t)}} \lesssim t^{\kappa/2} |\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}^{\rho,\eta}_{(0,t)}}, \tag{4.8}$$

where the implicit proportional constant is independent of $\boldsymbol{f}$ *and* t.

(3) *If* $\Pi^{\mathrm{g}}_x \tau$ *happens to be continuous for any* $\tau \in T$, *then the function*

$$\mathbf{R}^{\mathsf{M}} \boldsymbol{f} := \mathbf{R}^{\mathsf{M}}(C_t \boldsymbol{f})$$

satisfies

$$\mathbf{R}^{\mathsf{M}}\big(\mathcal{P}^{\mathsf{M}}_t \boldsymbol{f}\big)(x) = \int_{[0,x_0]\times\mathbb{T}^d} P(x, y) \mathbf{R}^{\mathsf{M}} \boldsymbol{f}(y) dy. \tag{4.9}$$

Proof. We provide only a sketch here. See [65, Theorem 5.9] for the detailed proof. Recall the decomposition $\mathbf{L}^{-1} = \mathbf{K} + \mathbf{K}'$. Noting that $\mathbf{R}^{\mathsf{M}} \boldsymbol{f} := \mathbf{R}^{\mathsf{M}}(C_t \boldsymbol{f}) \in \mathcal{C}^\eta$, we denote by $(\mathcal{K}')^{\mathsf{M}}$ the lift of the $\mathbf{K}'$ operator in the polynomial part of the regularity structure

$$\big((\mathcal{K}')^{\mathsf{M}} \boldsymbol{f}\big) := \sum_{|\ell|_{\mathfrak{s}} < \rho} \partial^\ell \mathbf{K}'\big(\mathbf{R}^{\mathsf{M}} \boldsymbol{f}\big) \frac{X^\ell}{\ell!} \in T_X.$$

Since $\mathbf{K}'$ maps $\mathcal{C}^\eta$ into $\mathcal{C}^\infty$, we can show that $(\mathcal{K}')^{\mathsf{M}} \boldsymbol{f} \in \mathcal{D}^{\rho,\eta+2}$. Then, we can also show that

$$\mathcal{P}^{\mathsf{M}}_t \boldsymbol{f} = \mathcal{Q}_{<\rho} \mathcal{K}^{\mathsf{M}}(C_t f) + (\mathcal{K}')^{\mathsf{M}} \boldsymbol{f} \in \mathcal{D}^{\rho,\eta+2}$$

by Theorem 17′. The property (1) is obvious from the definition. To show the property (3), note that

$$\mathbf{R}^{\mathsf{M}}\mathcal{P}_t^{\mathsf{M}}\boldsymbol{f} = \mathbf{K}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big) + \mathbf{K}'\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big) = \mathbf{L}^{-1}\big(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}\big).$$

Since $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}$ vanishes on $(-\infty,0)\times\mathbb{T}^d$ by Corollary 5, we have (4.9). To show the property (2), recall the sufficient condition for the equivalence between two norms $\|\!|\cdot\|\!|_{\mathcal{D}^{\gamma,\eta}}$ and $\|\!|\cdot\|\!|'_{\mathcal{D}^{\gamma,\eta}}$. Set $\boldsymbol{g} := \mathcal{P}_t^{\mathsf{M}}\boldsymbol{f}$. By definition, $\boldsymbol{g}$ takes values in the function-like sector V with $\alpha_0(V) = \beta_0 + 2$. If $\eta \le -1$, it is sufficient to show that

$$\lim_{t\downarrow 0} \boldsymbol{g}_{\mathbf{1}}(x) = 0$$

for the equivalence between $\|\!|\boldsymbol{g}\|\!|_{\mathcal{D}^{\rho,\eta+2}}$ and $\|\!|\boldsymbol{g}\|\!|'_{\mathcal{D}^{\rho,\eta+2}}$. Recall that

$$\boldsymbol{g}_{\mathbf{1}} = \mathbf{R}^{\mathsf{M}}\boldsymbol{g} = \mathbf{L}^{-1}(\mathbf{R}^{\mathsf{M}}\boldsymbol{f}).$$

Since $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} \in \mathcal{C}^{\eta}$, we have $\boldsymbol{g}_{\mathbf{1}} \in \mathcal{C}^{\eta+2}$ by Schauder estimate, so it is Hölder continuous. Since $\boldsymbol{g}_{\mathbf{1}} = 0$ on $(-\infty,0)\times\mathbb{T}^d$ from the non-anticipativity of $K_{\mathbf{L}}$, it also vanishes at $x_0 = 0$. If $-1 \le \eta \le 0$, we also have

$$\lim_{t\downarrow 0} \boldsymbol{g}_{X^k}(x) = 0$$

for any $|k|_{\mathfrak{s}} = 1$ by a similar argument. By the equivalence between $\|\!|\boldsymbol{g}\|\!|_{\mathcal{D}^{\rho,\eta+2}}$ and $\|\!|\boldsymbol{g}\|\!|'_{\mathcal{D}^{\rho,\eta+2}}$, we have (4.8) as follows:

$$\begin{aligned}\|\!|\boldsymbol{g}\|\!|_{\mathcal{D}^{\rho,\eta+2-\kappa}_{(0,t)}(0,T)} &\lesssim \|\!|\boldsymbol{g}\|\!|'_{\mathcal{D}^{\rho,\eta+2-\kappa}_{(0,t)}(0,T)} \lesssim t^{\frac{\kappa}{2}}\|\!|\boldsymbol{g}\|\!|'_{\mathcal{D}^{\rho,\eta+2}_{(0,t)}(0,T)}\\ &\lesssim t^{\frac{\kappa}{2}}\|\!|\boldsymbol{g}\|\!|_{\mathcal{D}^{\rho,\eta+2}_{(0,t)}(0,T)} \lesssim t^{\frac{\kappa}{2}}\|\!|\boldsymbol{f}\|\!|_{\mathcal{D}^{\gamma,\eta}_{(0,t)}(0,T)}.\end{aligned}$$ ■

4.4. Fixed-point solution

Finally, we make sense of the equation (4.2) and show its local well-posedness.

Definition. A *regularity structure* $\mathcal{T}$ is said to be *associated with equation* (4.1) if it satisfies Assumptions A–B and contains subcomodules S, DS, F, N of T satisfying Assumption A1 and the following constraints.

- The symbol Ξ and the set ∂S are contained in N.
- The sector S is function-like and regular products

$$S\times\cdots\times S \to F,\quad DS\times DS\to N,\quad F\times N\to T$$

  are given and satisfy Assumption A3. We denote them all by the same symbol $\star$.
- Abstract integration operators

$$\mathcal{I} : T\to S,\quad \mathcal{I}_{e_i} : T\to DS,\quad (1\le i\le d)$$

  are given and satisfy Assumption B.

- Derivative operators

$$D_i : S \to DS, \quad (1 \le i \le d)$$

are given and satisfy

$$\Pi \circ D_i = \partial_{x_i} \circ \Pi, \quad \text{and} \quad D_i X^k = k_i X^{k-e_i} \mathbf{1}_{k \ge e_i}, \quad \text{and} \quad D_i \mathcal{I}\tau = \mathcal{I}_{e_i}\tau.$$

The element Ξ represents the noise ζ. The spaces S and DS are used to represent the solution u and its derivative $\partial_{x'} u$, respectively. (The letter S is chosen for "solution".) The space F is used to represent $f(u)$ and $g_n(u)$, with $n = 0, 1, 2$. (The letter F is chosen for "function".) The space N is used to represent the "singular" elements ζ, $\partial_{x'} u$, and $(\partial_{x'} u)^2$. (The letter N is chosen for "noise".) The only role played by the intermediate spaces DS, F, N is to clarify on which spaces the product $\star$ is defined; they play no other role. We will see in Section 9 how to construct explicitly a regularity structure associated with the generalized (KPZ) equation. The $\star$ product is used to define nonlinear images of singular modelled distributions as in Section 2.4. In this setting, the regularity structure lift of the generalized (KPZ) equation is formulated under the form

$$\begin{aligned} \boldsymbol{v} &= \boldsymbol{h} + \mathcal{P}_t^{\mathsf{M}}\big(f^\star(\boldsymbol{v})\Xi + g^\star(\boldsymbol{v}, D\boldsymbol{v})\big) \\ &=: \Phi_t^{\boldsymbol{h},\mathsf{M}}(\boldsymbol{v}) \end{aligned} \tag{4.10}$$

for appropriate choices of $t \in (0, 1]$ and $\rho > 0$, where $f^\star(\boldsymbol{v})$ denotes the composition operator (Proposition 9), and $f^\star(\boldsymbol{v})\Xi := f^\star(\boldsymbol{v}) \star \Xi$, and

$$\begin{aligned} g^\star(\boldsymbol{v}, D\boldsymbol{v}) &:= \mathcal{Q}_{<\gamma'}\big\{g_2^\star(\boldsymbol{v}) \star (D\boldsymbol{v})^{\star 2} + g_1^\star(\boldsymbol{v}) \star (D\boldsymbol{v}) + g_0^\star(\boldsymbol{v})\big\} \\ &:= \mathcal{Q}_{<\gamma'}\left\{\sum_{i,j=1}^d (g_2^{ij})^\star(\boldsymbol{v}) \star (D_i\boldsymbol{v}) \star (D_j\boldsymbol{v}) + \sum_{i=1}^d (g_1^i)^\star(\boldsymbol{v}) \star (D_i\boldsymbol{v}) + g_0^\star(\boldsymbol{v})\right\} \end{aligned}$$

for an appropriate $\gamma' \in \mathbb{R}$ – we will choose $\gamma' = \gamma + \beta_0$ in the proof of Theorem 23.

Pick a $\mathbf{K}$-admissible model $\mathsf{M} = (g, \Pi)$ over $\mathcal{T}$ and $\boldsymbol{h} \in \mathcal{D}^{\gamma,\eta}(T_X, \mathrm{g})$. Assume that $\Phi^{\boldsymbol{h},\mathsf{M}}$ sends $\mathcal{D}^{\gamma,\eta}(S, \mathrm{g})$ into itself, which turns out to be the case as proved below under the conditions of Theorem 23.

Definition. A *solution to equation* (4.2) *on the time interval* $(0, t_0)$ is a fixed point of the map $\Phi_{t_0}^{\boldsymbol{h},\mathsf{M}} : \mathcal{D}_{(0,t_0)}^{\gamma,\eta}(S, \mathrm{g}) \to \mathcal{D}_{(0,t_0)}^{\gamma,\eta}(S, \mathrm{g})$.

Theorem 23. *Assume that f and g are smooth functions. Let $\mathcal{T}$ be a regularity structure associated with equation* (4.1) *and satisfying Assumptions* A–B, *with* $\beta_0 \in (-2, -1)$. *Pick* $\eta \in (0, \beta_0 + 2]$ *and* $\gamma > -\beta_0$. *Then, for any* $\mathbf{K}$*-admissible model* $\mathsf{M} = (g, \Pi)$ *and any* $\boldsymbol{h} \in \mathcal{D}^{\gamma,\eta}(T_X, \mathrm{g})$, *there exists a positive time* $t_0 = t_0(\boldsymbol{h}, \mathsf{M})$ *such that equation* (4.2) *has a unique solution* $\boldsymbol{u}$ *on the time interval* $(0, t_0)$. *The time* t_0 *can be chosen to be a lower semicontinuous function of* $\boldsymbol{h}$ *and* M.

Proof. Recall that $\mathcal{D}_\alpha^{\gamma,\eta}(T, \mathrm{g})$ denotes the set of modelled distributions of the form (4.7). Starting from $\boldsymbol{v} \in \mathcal{D}^{\gamma,\eta}(S, \mathrm{g})$, we show that

$$f^\star(\boldsymbol{v})\Xi + g^\star(\boldsymbol{v}, D\boldsymbol{v}) \in \mathcal{D}^{\gamma+\beta_0, 2\eta-2}(T, \mathrm{g}).$$

From the "singular/exploding" version of Proposition 9 given at the end of Section 4.2, one has $f^\star(\boldsymbol{v}), (g_2^{ij})^\star(\boldsymbol{v}), (g_1^i)^\star(\boldsymbol{v}), g_0^\star(\boldsymbol{v}) \in \mathcal{D}_0^{\gamma,\eta}(F,\mathrm{g})$. Since $\Xi \in \mathcal{D}_{\beta_0}^{\infty,\infty}(N,\mathrm{g})$, one has

$$f^\star(\boldsymbol{v})\Xi \in \mathcal{D}^{\gamma+\beta_0,\eta+\beta_0}(T,\mathrm{g}) \subset \mathcal{D}^{\gamma+\beta_0,2\eta-2}(T,\mathrm{g})$$

from the singular version of Proposition 8 given at the end of Section 4.2. Noting that the smallest homogeneity in the subcomodule ∂S is $\beta_0 + 1 < 0$, which is the homogeneity of $\mathcal{I}_{e_i}\Xi$, one has $D_i\boldsymbol{v} \in \mathcal{D}_{\beta_0+1}^{\gamma-1,\eta-1}(\partial S,\mathrm{g})$ and $(D_i\boldsymbol{v}) \star (D_j\boldsymbol{v}) \in \mathcal{D}_{2\beta_0+2}^{\gamma+\beta_0,2\eta-2}(N,\mathrm{g})$. Thus, $g^\star(\boldsymbol{v}, D\boldsymbol{v}) \in \mathcal{D}^{\gamma+\beta_0,2\eta-2}(T,\mathrm{g})$. From Proposition 22, one has

$$\begin{aligned}\|\!|\Phi_t^{\boldsymbol{h},\mathsf{M}}(v)\|\!|_{\mathcal{D}^{\gamma,\eta}_{(0,t)}} &\lesssim \|\!|\boldsymbol{h}\|\!|_{\mathcal{D}^{\gamma,\eta}_{(0,t)}} + t^{\eta/2}\|\!| f^\star(\boldsymbol{v})\Xi + g^\star(\boldsymbol{v}, D\boldsymbol{v})\|\!|_{\mathcal{D}^{\gamma+\beta_0,2\eta-2}_{(0,t)}} \\ &\lesssim \|\!|\boldsymbol{h}\|\!|_{\mathcal{D}^{\gamma,\eta}_{(0,t)}} + t^{\eta/2} C\left(\|\!|\boldsymbol{v}\|\!|_{\mathcal{D}^{\gamma,\eta}_{(0,t)}}\right)\end{aligned}$$

for some locally bounded function C. Then, one can associate with each positive radius $\lambda \gtrsim \|\!|\boldsymbol{h}\|\!|_{\mathcal{D}^{\gamma,\eta}}$ a time horizon $t(\lambda)$ such that $\Phi_{t(\lambda)}^{\boldsymbol{h},\mathsf{M}}$ sends the ball of $\mathcal{D}^{\gamma,\eta}_{(0,t(\lambda))}(S,\mathrm{g})$ of radius λ into itself. From the local Lipschitz continuity result, the map $\Phi_{t(\lambda)}^{\boldsymbol{h},\mathsf{M}}$ is also a contraction on the ball of $\mathcal{D}^{\gamma,\eta}_{(0,t(\lambda))}(S,\mathrm{g})$ of radius λ. As such, it has a unique fixed point on the ball of radius λ. An elementary argument gives the uniqueness of a fixed point within $\mathcal{D}^{\gamma,\eta}_{(0,t(\lambda))}(S,\mathrm{g})$, as in the proof of [57, Theorem 4.7]. ■

The proof makes it clear that one can ask f and g to have finite regularity rather than being smooth. We do not try to optimize the regularity assumptions on f and g here. Thinking of $\boldsymbol{h}$ as the regularity structure lift of the free propagation of an initial condition on $\mathbb{T}^d$, assuming in $\boldsymbol{h} \in \mathcal{D}^{\gamma,\eta}(T_X,\mathrm{g})$ allows us to work with an initial condition of Hölder regularity η – recall the constraint $\eta \in (0, \beta_0 + 2]$. Note that the map is uniformly contracting on a small enough time interval for f and g ranging in a bounded set. In order to compare fixed points of $\Phi^{\boldsymbol{h},\mathsf{M}}$ associated with different admissible models over $\mathcal{T}$ – hence different maps on different spaces, we use the metric d_γ in Definition 3 with a slight modification to $\mathcal{D}^{\gamma,\eta}_{(0,t)}$ norms. One can then prove the following statement in terms of this metric by making explicit in the reconstruction theorem and the lifting theorem that the operators $\mathbf{R}^\mathsf{M}$ and $\mathscr{P}_t^\mathsf{M}$ depend in a locally Lipschitz way on M with respect to the pseudo-distance d_γ on the space of models over $\mathcal{T}$ introduced in (2.21). We do not give the details here and refer the reader to the corresponding results in [53], Theorem 3.10 and Theorem 5.12 therein.

Proposition 24. *Given any time $t_0' < t_0(\boldsymbol{h},\mathsf{M})$, the restriction to $(0,t_0'] \times \mathbb{T}^d$ of $\boldsymbol{u}$ defines locally a continuous function of $\boldsymbol{h} \in \mathcal{D}^{\gamma,\eta}(T_X,\mathrm{g})$ and the* **K***-admissible model* M.

Together with Theorem 43 in Section 6.3 below and Chandra and Hairer's convergence result [26], this continuity result allows to give meaning of the solution to a singular stochastic PDE as a limit in probability of solutions of renormalized equations driven by a mollified noise. This result holds more generally for all the equations that can be treated using regularity structures. Emphasize that this continuity result is fundamental. In a random setting where the noise is random and the models of interest are constructed as measurable functionals of the noise, the continuity allows to transport automatically support theorems or large deviation results about random models into corresponding results about the solutions of the regularity structure lifts of the equations under study. See Hairer and Schönbauer's work [60] on support theorems, Hairer and Weber's work [62] on large deviation results, or Hairer and Mattingly's work [59] on the strong Feller property for solutions of singular stochastic PDEs, for a sample.

The last statement of this section makes the link between solving equation (4.1) with a smooth noise ζ and the corresponding problem in the regularity structure equipped with the canonical model M^ζ associated with the smooth noise. The latter is constructed in Section 6.1, and the only thing we presently need to know about it is that its reconstruction map $\mathbf{R}^{\mathsf{M}^\zeta}$ is multiplicative with respect to the $\star$-product of modelled distributions and sends the noise symbol Ξ on the smooth function ζ. For positive exponents $\gamma \in (-\beta_0, 2)$ and $\eta \in (0, \beta_0 + 2]$, pick $v \in \mathcal{C}^\eta(\mathbb{T}^d)$ and denote by $P_\gamma v$ the lift in the polynomial structure of the heat propagator acting on v, defined by (4.5).

Proposition 25. *Let $\boldsymbol{u} \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g})$ stand for the solution in a sufficiently small time interval $(0, t_0)$ of the fixed-point problem*

$$\boldsymbol{u} = P_\gamma v + \mathcal{P}^{\mathsf{M}^\zeta}_{t_0}\big(f^\star(\boldsymbol{u})\,\Xi + g^\star\big(\boldsymbol{u}, D\boldsymbol{u}\big)\big). \tag{4.11}$$

Then, on the domain $(0, \frac{t_0}{2}) \times \mathbb{T}^d$, the function

$$u := \mathbf{R}^{\mathsf{M}^\zeta}(C_{t_0}\boldsymbol{u})$$

coincides with the solution to the well-posed equation (4.1) with initial condition v.

Proof. As in (4.9), the function u satisfies the equation

$$u(x) = Pv(x) + \int_{(0,x_0)\times\mathbb{T}^d} P(x,y)\mathbf{R}^{\mathsf{M}^\zeta} C_{t_0}\big(f^\star(\boldsymbol{u})\Xi + g^\star\big(\boldsymbol{u}, D\boldsymbol{u}\big)\big)(y)dy,$$

with Pv the free propagation of the initial condition. We take advantage of the fact that M^ζ is a smooth model to write

$$\mathbf{R}^{\mathsf{M}^\zeta}(\boldsymbol{w})(x) = \Pi^\zeta_x\big(\boldsymbol{w}(x)\big)(x)$$

for any modelled distribution $\boldsymbol{w} \in \mathcal{D}^{\alpha,\eta}(T, \mathrm{g}^\zeta)$ with $\alpha > 0$ – see identity (2.32). Moreover, for any $\boldsymbol{f} \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g}^\zeta)$, we have $C_{t_0}\boldsymbol{f}(x) = \boldsymbol{f}(x)$ if $x \in (0, \frac{t_0}{2}) \times \mathbb{T}^d$. Thus, in this domain, we can use the multiplicative character of the map $\mathbf{R}^{\mathsf{M}^\zeta}$ and write

$$\mathbf{R}^{\mathsf{M}^\zeta}\big(f(\boldsymbol{u})\big) = f\big(\mathbf{R}^{\mathsf{M}^\zeta}\boldsymbol{u}\big), \quad \mathbf{R}^{\mathsf{M}^\zeta}(D\boldsymbol{u}) = \partial_x \mathbf{R}^{\mathsf{M}^\zeta}\boldsymbol{u}$$

as a consequence of Corollary 7 and Proposition 10, and

$$\mathbf{R}^{\mathsf{M}^\zeta}\big(f(\boldsymbol{u})\Xi + g\big(\boldsymbol{u}, D\boldsymbol{u}\big)\big) = f\big(\mathbf{R}^{\mathsf{M}^\zeta}\boldsymbol{u}\big)\zeta + g\big(\mathbf{R}^{\mathsf{M}^\zeta}\boldsymbol{u}, \partial_{x'}\mathbf{R}^{\mathsf{M}^\zeta}\boldsymbol{u}\big).$$

This finishes the proof. ■

Arrived at that stage, we have a model-dependent notion of solution u^{M} to the generalized (KPZ) equation, under the form

$$u^{\mathsf{M}} = \mathbf{R}^{\mathsf{M}}(\boldsymbol{u}^{\mathsf{M}}), \quad \boldsymbol{u}^{\mathsf{M}} = \Phi^{\boldsymbol{h},\mathsf{M}}(\boldsymbol{u}),$$

indexed by the set of **K**-admissible models M on the regularity structure $\mathcal{T}$ associated with the equation. Theorem 19 gives us a whole family of smooth **K**-admissible models which we can use. However, the **K**-admissible models of interest are not smooth as we wish they satisfy the identity $\Pi\Xi = \zeta$ for a non-smooth noise ζ. The combinatorial structure of the elements of $\mathcal{T}$ detailed in Section 9 allows to associate to any regularized version ζ_ε of ζ a **K**-admissible model Π^ε such that $\Pi^\varepsilon\Xi = \zeta_\varepsilon$, and Π^ε is multiplicative for the $\star$-product on T. We will talk of Π^ε as the naive interpretation map. However, these models diverge as $\varepsilon > 0$ goes to 0. The tools needed to construct some ε-dependent smooth models that have a limit as ε goes to 0 are developed in the next section at the same level of generality as Section 2 and Section 3. The so-called renormalization operation involved in the construction of these converging **K**-admissible models will be given a dynamical meaning in Section 6. It is only after Theorem 43 in Section 6 that we will be able to give an answer to the question "*What dynamics does* u^{M} *follow?*" when M is the renormalized naive model associated with the Π-map ${}^{k_\varepsilon}\Pi^\varepsilon$ introduced in the next section.

5. Renormalization structures

We introduce in this section the fundamental notion of renormalization structure and a notion of compatibility between some regularity and renormalization structures. We emphasized in the previous paragraph that it is generically not possible to define a canonical **K**-admissible model as a limit of the canonical **K**-admissible models associated with regularized noises ζ_ε if the noise(s) ζ is (are) not sufficiently regular. On a technical level, the non-convergence of the models M^ε is related to the fact that the canonical model is defined by some intricate convolution of kernels that explode on the diagonal. Limit models need to be constructed by probabilistic means as limits in probability of models built from regularized noises, *using a moving window*, as in Meta-theorem 1 in Section 1. The implementation of this moving window picture involves the renormalization structures that we introduce in this section. Note that we do not need to know the details of the renormalization operation; the only properties that we need are encoded in the definition of a renormalization structure and the compatibility condition with a regularity structure given below in Definition 26. An example of renormalization structure will be given in Section 9, where the renormalization operation will be intimately related to the Taylor expansion procedure.

Renormalization structures are defined in Section 5.1. If we call the concrete regularity structures from Section 2.2 right regularity structures, then renormalization structures

$$\mathcal{U} = \big((U, \delta), (U^-, \delta^-)\big)$$

look like left regularity structures, with the difference that elements of the space U^- have non-positive homogeneities. A fundamental notion of compatibility between renormalization and regularity structures is introduced in Section 5.2; it accounts for the fact that the renormalization operation induces a renormalization operation on T^+ and "commutes" with the recentering operators Δ and Δ^+. This property allows to associate with each model M over $\mathcal{T}$ and each character k on U^- a new model ${}^k\mathsf{M}$ on $\mathcal{T}$. This is the main result of Section 5.2, Theorem 28. A large class of characters k produces some $\mathbf{K}$-admissible models ${}^k\mathsf{M}$ if M is $\mathbf{K}$-admissible.

5.1. Definition

A renormalization structure is made up of two ingredients. First, it is a vector space U with a basis whose elements are built by induction from elementary elements and multilinear operators giving new elements. The use of the symbol τ for a generic basis vector emphasizes this recursive, tree-like, definition. Each basis vector τ is a placeholder for a function $[\tau]$ from $(0,1]$ into a Banach space, typically $\mathbb{R}$, $\mathbb{C}$, a Hölder space or an algebra, whose structure as an element of the target space is encoded in the structure of τ. In the cases of interest, the functions $[\tau]$ have no limit in 0^+ and the basic problem is to remove in a "consistent" way the diverging pieces of these $[\tau]$ so as to end up with a collection of functions parametrized by $\varepsilon > 0$ having a limit where ε goes to 0. The functions $[\tau]$ are then said to have been renormalized. What "consistent" means is part of what follows.

Roughly speaking, the basic operation for renormalizing a placeholder τ consists in removing from τ its different diverging pieces, in all possible sensible ways. This is the second ingredient of a renormalization structure. Tuples of pieces of elements of U are not necessarily elements of U; we store them in a side space U^-. Endowing U^- with an algebra structure allows to store the removed pieces of τ as an element of U^- under the form of a product. We require nonetheless that any τ amputated from diverging pieces is an element of U; this is a restriction on which pieces of any $\tau \in U$ can be removed. We thus have a splitting map

$$\delta : U \to U^- \otimes U,$$

with $\delta\tau$ the sum of all the elements from $U^- \otimes U$ corresponding to removing from τ all possible diverging allowed pieces, possibly several at a time. The removed pieces may themselves have diverging subpieces, and it makes sense to assume that we have another splitting map

$$\delta^- : U^- \to U^- \otimes U^-,$$

which extracts them on the left-hand side of the tensor product $U^- \otimes U^-$. The fact that the remaining piece is still in U^- rather than in another space is a consistency requirement.

Definition 26. A *renormalization structure* is a pair of graded vector spaces

$$U =: \bigoplus_{\beta \in B} U_\beta, \quad U^- =: \bigoplus_{\alpha \in B^-} U^-_\alpha$$

such that the following holds.

- The vector spaces U_α^- and U_β are finite dimensional.
- The space U^- is a connected graded bialgebra with unit $\mathbf{1}_-$, counit $\mathbf{1}'_-$, coproduct

$$\delta^- : U^- \to U^- \otimes U^-,$$

  and grading $B^- \subset (-\infty, 0]$, with $0 \in B^-$.
- The index set B for U is a locally finite subset of $\mathbb{R}$ bounded below. The space U is a left comodule over U^-; that is, U is equipped with a splitting map $\delta : U \to U^- \otimes U$, which satisfies

$$(\mathrm{Id} \otimes \delta)\delta = (\delta^- \otimes \mathrm{Id})\delta \quad \text{and} \quad (\mathbf{1}'_- \otimes \mathrm{Id})\delta = \mathrm{Id}. \tag{5.1}$$

  Moreover, for any $\beta \in B$, one has

$$\delta U_\beta \subset \bigoplus_{\alpha \leq 0} U_\alpha^- \otimes U_{\beta-\alpha}. \tag{5.2}$$

We denote by

$$\mathcal{U} := \big((U, \delta), (U^-, \delta^-)\big)$$

a renormalization structure.

Similarly to the regularity structure, let $\mathcal{U}_\alpha^-$ and $\mathcal{U}_\beta$ be bases of U_α^- and U_β, respectively, and set

$$\mathcal{U}^- := \bigcup_{\alpha \in B^-} \mathcal{U}_\alpha^-, \quad \mathcal{U} := \bigcup_{\beta \in B} \mathcal{U}_\beta.$$

Note that, unlike in the definition of a concrete regularity structure satisfying Assumption A1, we do not require that $\mathcal{U}_0$ is one dimensional in the definition of a renormalization structure. Since all $\alpha \in B^-$ are non-positive, one has $\beta - \alpha \geq \beta$ in (5.2). Proposition 48 in Appendix B can be applied to the negative grading B^- of U^-, and says that U^- is a Hopf algebra; we denote by S_- its antipode. Choosing a basis of U^- provides an associated decomposition of $\delta\tau$ of the form

$$\delta\tau =: \sum_{\varphi \trianglelefteq \tau} \varphi \otimes \tau/^- \varphi,$$

where the φ are distinct elements of the chosen basis. The notation $\varphi \trianglelefteq \tau$ means that φ is a basis element that appears as one of the left-hand side members of the finite sum giving $\delta\tau$. We call δ a *renormalization splitting* and fix throughout a basis of U^-. The results we prove in the sequel do not depend on that arbitrary choice. Similarly to what we saw in Section 2.2 for the Hopf algebra (T^+, Δ^+), the δ^- splitting of the Hopf algebra (U^-, δ^-) induces a convolution group law on the set G^- of characters on U^-

$$(k_1 * k_2)\tau := (k_1 \otimes k_2)\delta^-\tau, \quad (\tau \in U^-).$$

The inverse of a character k for the convolution product is explicit and given by $k \circ S_{-}$. Given a character k on U^{-}, we define a linear map $\tilde{k} : U \to U$, setting

$$\tilde{k} := (k \otimes \mathrm{Id})\delta.$$

The group G^{-} acts on U from right. Indeed, as a direct consequence of the comodule property in (5.1), one has

$$\begin{aligned}\widetilde{k_1 * k_2} &= ((k_1 * k_2) \otimes \mathrm{Id})\delta = (k_1 \otimes k_2 \otimes \mathrm{Id})(\delta^{-} \otimes \mathrm{Id})\delta \\ &= (k_1 \otimes k_2 \otimes \mathrm{Id})(\mathrm{Id} \otimes \delta)\delta = (k_1 \otimes \widetilde{k_2})\delta \\ &= \widetilde{k_2} \circ \widetilde{k_1}\end{aligned}$$

for any $k_1, k_2 \in G^{-}$.

5.2. Compatible renormalization and regularity structures

We introduce a "compatibility" property between regularity and renormalization structures. We use the notations from Appendix B. In particular, given an algebra A and two spaces E, F, we define a linear map $\mathcal{M}^{(13)}$ from the algebraic tensor product $A \otimes E \otimes A \otimes F$ to the algebraic tensor product $A \otimes E \otimes F$ setting

$$\mathcal{M}^{(13)}(a_1 \otimes e \otimes a_2 \otimes f) := (a_1 a_2) \otimes e \otimes f.$$

Recall we write $\mathcal{T} = ((T^{+}, \Delta^{+}), (T, \Delta))$ for a regularity structure and S_{+} for the antipode map on T^{+}.

Definition 27. A regularity structure $\mathcal{T}$ is said to be *compatible* with a renormalization structure $\mathcal{U}$ if the following three compatibility conditions hold true.

(a) The spaces T and U coincide as linear spaces and the bases $\mathcal{B}$ and $\mathcal{U}$ coincide. (Each element $\tau \in \mathcal{B}$ is in particular homogeneous in both T and U, but it may belong to $\mathcal{B}_{\beta_1}$ and $\mathcal{U}_{\beta_2}$ with $\beta_1 \neq \beta_2$.) Moreover,

$$\delta T_{\beta} \subset U^{-} \otimes T_{\beta} \quad \text{for all } \beta \in A. \tag{5.3}$$

(b) There exists an algebra morphism

$$\delta^{+} : T^{+} \to U^{-} \otimes T^{+}$$

such that

$$\big(\mathrm{Id} \otimes \delta^{+}\big)\delta^{+} = \big(\delta^{-} \otimes \mathrm{Id}\big)\delta^{+} \quad \text{and} \quad \big(\mathbf{1}'_{-} \otimes \mathrm{Id}\big)\delta^{+} = \mathrm{Id} \tag{5.4}$$

and

$$\delta^{+} T^{+}_{\alpha} \subset U^{-} \otimes T^{+}_{\alpha} \quad \text{for all } \alpha \in A^{+}. \tag{5.5}$$

(c) The compatibility conditions

$$\big(\mathrm{Id}\otimes\Delta^{(+)}\big)\delta^{(+)} = \mathcal{M}^{(13)}\big(\delta^{(+)}\otimes\delta^{+}\big)\Delta^{(+)} \tag{5.6}$$

and

$$\big(\mathrm{Id}\otimes\mathbf{1}'_{+}\big)\delta^{+} = \mathbf{1}'_{+}(\cdot)\mathbf{1}_{-} \tag{5.7}$$

hold.

Emphasize the fact that the homogeneity notion in T captures the notion of regularity of the associated analytic objects encoded by elements of T while the homogeneity notion in U captures the diverging behaviour of the corresponding regularized objects, as the regularization parameter goes to 0. It makes sense that the two notions of homogeneities are unrelated. Definition 27 also captures the fact that the renormalization procedure encoded in $\mathcal{U}$ induces a renormalization operation on T^+ and commutes with the recentering operators Δ and Δ^+. We will see in Proposition 30 below that the six conditions from Definition 27 hold iff condition (5.3) and condition (5.6), in its form without the $+$ labels, hold, under a reasonable assumption on δ^+ that holds true for the regularity and renormalization structures associated with (systems of) singular stochastic PDEs.

Compare conditions (5.3) and (5.2). Emphasize here as in item (a) that the notion of homogeneity is relative to the grading used to define it. An element of $T = U$ may thus have different homogeneities, depending on whether it is considered as an element of T or U. By condition (a), the space T is a left U^--comodule. The map δ^+ in (b) accounts for the effect in T^+ of the renormalization process. By (5.4), the space T^+ is also a left U^--comodule. Hence, for given a character k on U^-, we can define linear maps $\tilde{k} : T \to T$ and $\tilde{k}^+ : T^+ \to T^+$ by

$$\tilde{k} = \big(k\otimes\mathrm{Id}\big)\delta, \quad \text{and} \quad \tilde{k}^+ = \big(k\otimes\mathrm{Id}\big)\delta^+.$$

Properties (5.3) and (5.5) ensure that homogeneities of elements of T and T^+ are stable under these actions. Condition (c), read with the $+$ labels, somehow says that the renormalization operation encoded in $\tilde{k}$ commutes with the Taylor expansion operation on the coefficients of any modelled distribution, encoded in Δ^+. Condition (c), read without the $+$ labels, says something similar for modelled distributions. Note that the Hopf algebra T^+ is a left U^--comodule bialgebra. By Proposition 49, we have the following compatibility condition on the antipode:

$$\delta^+\circ S_+ = \big(\mathrm{Id}\otimes S_+\big)\circ\delta^+. \tag{5.8}$$

Recall that given a model $\mathsf{M} = (g, \Pi)$ on $\mathcal{T}$, the anchored interpretation operator Π_x^{g} associated with M is given for any $x\in\mathbb{R}^d$ by

$$\Pi_x^{\mathrm{g}} = (\Pi\otimes \mathrm{g}_x^{-1})\Delta.$$

The next statement and its proof are part of Theorem 6.15 in Bruned, Hairer, and Zambotti's work [20] on the algebraic renormalization of regularity structures. It tells us that the $\tilde{k}$ and $\tilde{k}^+$ maps have jointly a natural and simple action on the space of models on $\mathcal{T}$.

Theorem 28. *Let a renormalization structure* $\mathcal{U} = (U, U^-)$ *be compatible with a regularity structure* $\mathcal{T} = (T^+, T)$*. Given any character* k *on* U^-*, and any model* $\mathsf{M} = (g, \Pi)$ *on* $\mathcal{T}$*, define* ${}^k\mathsf{M} = ({}^k\mathrm{g}, {}^k\Pi)$*, on* $\mathcal{T}$ *setting*

$$ {}^k\mathsf{M} := \big(\mathrm{g} \circ \tilde{k}^+, \Pi \circ \tilde{k}\big). $$

One has

$$ \big(\mathrm{g}_y \circ \tilde{k}^+\big) * \big(\mathrm{g}_x \circ \tilde{k}^+\big)^{-1} = \mathrm{g}_{yx} \circ \tilde{k}^+ \tag{5.9} $$

and

$$ \big((\Pi \circ \tilde{k}) \otimes \big(\mathrm{g}_x \circ \tilde{k}^+\big)^{-1}\big)\Delta = \Pi_x^{\mathrm{g}} \circ \tilde{k} \tag{5.10} $$

for any $x, y \in \mathbb{R}^d$*. Moreover, the size conditions* (2.19) *and* (2.20) *hold for* ${}^k\mathsf{M} = ({}^k\mathrm{g}, {}^k\Pi)$*, so* ${}^k\mathsf{M}$ *is a model.*

Proof. The proof is short and simple because the notion of compatibility between some regularity and renormalization structures is tailored for that purpose. One has

$$ \begin{aligned} \big(\mathrm{g}_y \circ \tilde{k}^+\big) * \big(\mathrm{g}_x \circ \tilde{k}^+\big)^{-1} &= \big((k \otimes \mathrm{g}_y)\delta^+\big) \otimes \big((k \otimes \mathrm{g}_x)\delta^+ \circ S_+\big)\Delta^+ \\ &\overset{(5.8)}{=} \big((k \otimes \mathrm{g}_y)\delta^+\big) \otimes \big((k \otimes \mathrm{g}_x^{-1})\delta^+\big)\Delta^+ \\ &= \big(\mathrm{g}_y \otimes \mathrm{g}_x^{-1}\big) \circ \big(\tilde{k}^+ \otimes \tilde{k}^+\big)\Delta^+ \\ &\overset{(5.6)}{=} \big(\mathrm{g}_y \otimes \mathrm{g}_x^{-1}\big) \circ (k \otimes \Delta^+)\delta^+ \\ &= \mathrm{g}_{yx} \circ \tilde{k}^+ \end{aligned} $$

and

$$ \begin{aligned} \big((\Pi \circ \tilde{k}) \otimes \big(\mathrm{g}_x \circ \tilde{k}^+\big)^{-1}\big)\Delta &\overset{(5.8)}{=} \big((k \otimes \Pi)\delta \otimes (k \otimes \mathrm{g}_x^{-1})\delta^+\big)\Delta \\ &= \big(\Pi \otimes \mathrm{g}_x^{-1}\big) \circ \big(\tilde{k} \otimes \tilde{k}^+\big)\Delta \\ &\overset{(5.6)}{=} \big(\Pi \otimes \mathrm{g}_x^{-1}\big) \circ (k \otimes \Delta)\delta \\ &= \Pi_x^{\mathrm{g}} \circ \tilde{k}. \end{aligned} $$

The size conditions (2.19) and (2.20) on ${}^k\mathsf{M}$ follow now from formulas (5.9) and (5.10), and from the fact that the maps $\tilde{k}$ and $\tilde{k}^+$ preserve the spaces T_β and T_α^+, respectively, as a consequence of the stability conditions (5.3) and (5.5). ■

Together with Corollary 6, this statement implies in particular that if the model M takes values in the space of *continuous functions*, then the reconstruction operator ${}^k\mathbf{R}$ associated with the renormalized model is related to the reconstruction operator $\mathbf{R}$ associated with the unrenormalized model by the relation

$$ {}^k\mathbf{R} = \mathbf{R} \circ \tilde{k}. $$

This point will be used crucially in Section 6.3, where we will give a dynamical picture of the renormalization of models.

We consider in the remainder of this section the case of interest for the study of (systems of) singular stochastic PDE(s) where the regularity structure $\mathcal{T}$ is built from integration operators and satisfies Assumption B. Unfortunately, even if a model M is **K**-admissible, ${}^{k}\mathsf{M}$ is not always **K**-admissible for any $k \in G^-$. We put forward an assumption under which one builds **K**-admissible models using elements k of a non-trivial subgroup G^-_{ad} of G^-. Assume $\mathcal{B} = \mathcal{U}$ and let $\mathcal{F}$ stand for the family of operators

$$\mathcal{F} := \{\mathcal{I}_p\}_{|p|_{\mathfrak{s}} \le 1} \cup \{X^n \star\}_{n \in \mathbb{N}^{d+1}\setminus\{0\}},$$

acting on the basis $\mathcal{B} = \mathcal{U}$, where $X^n\star$ denotes the linear operator on T defined by

$$\tau \mapsto X^n \star \tau.$$

Recall that such a multiplication is always given by Assumption A3.

Assumption C1. The regularity structure $\mathcal{T}$ is built from integration operators and satisfies Assumption B1 and the renormalization structure $\mathcal{U}$ is compatible with $\mathcal{T}$. Moreover, the following holds.

- The algebra U^- is generated by the basis elements $\mathcal{U}_{<0} := \bigcup_{\alpha<0} \mathcal{U}_\alpha$ and the unit $\mathbf{1}_-$.
- Let $\mathfrak{J}^-$ be the ideal of U^- generated by the set $(\mathcal{F}(\mathcal{U})) \cap \mathcal{U}_{<0}$. The linear map $\delta : U \to U^- \otimes U$ satisfies, for any operator $F \in \mathcal{F}$ and $\tau \in \mathcal{U}$,
$$\delta(F\tau) - (\mathrm{Id} \otimes F)\delta\tau \in \mathfrak{J}^- \otimes U. \tag{5.11}$$
- We define a projection operator $P_- : U \to U^-$ setting $P_-\tau := \tau \mathbf{1}_{\tau \in \mathcal{U}_{<0}}$ for any $\tau \in \mathcal{U}$. The linear map
$$\delta^- : U^- \to U^- \otimes U^-$$
is defined by $\delta^- = (\mathrm{Id} \otimes P_-)\delta$ on $\mathcal{U}_{<0}$ and its multiplicative extension.

  Define the subset G^-_{ad} of G^- by
$$G^-_{\mathrm{ad}} := \big\{k \in G^-;\ k\big(F(\tau)\big) = 0 \text{ for any } F(\tau) \in \big(\mathcal{F}(\mathcal{U})\big) \cap \mathcal{U}_{<0}\big\}.$$

Proposition 29. *The set G^-_{ad} is a subgroup of G^-, and for any $k \in G^-_{\mathrm{ad}}$ and any* **K**-*admissible model M, one has ${}^{k}\mathsf{M}$ is also* **K**-*admissible. The group G^-_{ad} is called the renormalization group.*

The definition of the group G^-_{ad} gives the meaning to assumption (5.11). Up to irrelevant terms for $k \in G^-_{\mathrm{ad}}$, the renormalization operations in U^- or U of an "integral" is the integral of its renormalized integrand, and multiplication by a polynomial has no effect on the renormalization process.

Proof. Note that $k(\mathfrak{J}^-) = 0$ for any $k \in G^-_{\mathrm{ad}}$. Let τ be an element of $\mathcal{U}_\alpha$ such that $F\tau \in \mathcal{U}_{<0}$ for some $F \in \mathcal{F}$.

(a) Given $k, h \in G^-_{\mathrm{ad}}$ since identity (5.11) and the third property of Assumption C1 ensures that

$$(k * h)(F\tau) = (k \otimes h)\delta^-(F\tau) = 0,$$

for all $F \in \mathcal{F}$, we have $k * h \in G^-_{\mathrm{ad}}$. Next, we show that

$$k^{-1} = k \circ S_- \in G^-_{\mathrm{ad}}.$$

Denote by $\mathcal{M}^-$ the multiplication operator in U^- and pick $\sigma \in U_\beta$ with $\beta < 0$. Since $\delta\sigma \in \mathbf{1}_- \otimes \sigma + \sum_{\alpha<0} U^-_\alpha \otimes U_{\beta-\alpha}$, by applying the operator $\mathcal{M}^-(\mathrm{Id} \otimes S_- P_-)$ to (5.11), we have from Assumption B1 and the fact that $\mathfrak{J}^-$ is an ideal

$$S_-(F\sigma) \in \sum_{\alpha<0} \mathcal{M}^-\big(U^-_\alpha \otimes S_-\big(P_- F U_{\beta-\alpha}\big)\big) + \mathfrak{J}^-,$$

which implies $k^{-1}(FU_\beta) = k(S_- F U_\beta) = 0$, by an induction on β.

(b) Let $F = \mathcal{I}_p$. By (5.11),

$${}^k\Pi(\mathcal{I}_p\tau) = (k \otimes \Pi)\delta\mathcal{I}_p\tau = (k \otimes \Pi\mathcal{I}_p)\delta\tau = \partial^p\mathbf{K}(k \otimes \Pi)\delta\tau = \partial^p\mathbf{K}({}^k\Pi\tau).$$

Since we have a similar identity for $F = X^n\star$, we obtain that ${}^k\mathsf{M}$ is admissible.

■

We end this section by showing that the definition of compatible renormalization and regularity structures takes then a simple form under the following additional mild assumption. It essentially says that multiplications by a polynomial and integrations are not the sources of renormalization problems.

Assumption C2. The algebra morphism $\delta^+ : T^+ \to U^- \otimes T^+$ is determined by the identities

$$\delta^+ X^\ell_+ = \mathbf{1}_- \otimes X^\ell_+, \quad \delta^+(\mathcal{I}^+_p\tau) = (\mathrm{Id} \otimes \mathcal{I}^+_p)\delta\tau. \tag{5.12}$$

Proposition 30. *Under Assumption C2, assume that $\mathcal{T}$ satisfies the property (5.3) and the version of identity (5.6) without the $+$ labels. Then, the other conditions in Definition 27 follow automatically.*

Proof. The comodule property (5.4) follows from (5.1) and the definition (5.12). Indeed,

$$\begin{aligned}\big(\mathrm{Id} \otimes \delta^+\big)\delta^+(\mathcal{I}^+_n\tau) &= \big(\mathrm{Id} \otimes \delta^+\mathcal{I}^+_n\big)\delta\tau = \big(\mathrm{Id} \otimes \mathrm{Id} \otimes \mathcal{I}^+_n\big)(\mathrm{Id} \otimes \delta)\delta\tau \\ &= \big(\mathrm{Id} \otimes \mathrm{Id} \otimes \mathcal{I}^+_n\big)(\delta^- \otimes \mathrm{Id})\delta\tau = \big(\delta^- \otimes \mathrm{Id}\big)\big(\mathrm{Id} \otimes \mathcal{I}^+_n\big)\delta\tau \\ &= \big(\delta^- \otimes \mathrm{Id}\big)\delta^+(\mathcal{I}^+_n\tau).\end{aligned}$$

The counit parts of (5.4) and (5.7) are left to readers. The condition (5.5) follows from (5.3) and the definition (5.12). The $(+)$-labelled version of (5.6) is checked for $\mathcal{I}^+_n\tau \in \mathcal{B}^+$ as follows:

$$\begin{aligned}
&\mathcal{M}^{(13)}\big(\delta^+ \otimes \delta^+\big)\Delta^+(\mathcal{I}_n^+ \tau) \\
&\quad = \mathcal{M}^{(13)}\Big(\big(\delta^+ \mathcal{I}_n^+ \otimes \delta^+\big)\Delta\tau + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d} \delta^+ \frac{X_+^\ell}{\ell!} \otimes \delta^+(\mathcal{I}_{n+\ell}^+ \tau)\Big) \\
&\quad = \mathcal{M}^{(13)}\Big(\big((\mathrm{Id} \otimes \mathcal{I}_n^+)\delta \otimes \delta^+\big)\Delta\tau + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d} \mathbf{1}_- \otimes \frac{X_+^\ell}{\ell!} \otimes (\mathrm{Id} \otimes \mathcal{I}_{n+\ell}^+)\delta\tau\Big) \\
&\quad = \big(\mathrm{Id} \otimes \mathcal{I}_n^+ \otimes \mathrm{Id}\big)\mathcal{M}^{(13)}(\delta \otimes \delta^+)\Delta\tau + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d, \varphi \trianglelefteq \tau} \varphi \otimes \frac{X_+^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+(\tau/^-\varphi) \\
&\quad = \big(\mathrm{Id} \otimes \mathcal{I}_n^+ \otimes \mathrm{Id}\big)(\mathrm{Id} \otimes \Delta)\delta\tau + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d, \varphi \trianglelefteq \tau} \varphi \otimes \frac{X_+^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+(\tau/^-\varphi)
\end{aligned}$$

and

$$\begin{aligned}
\big(\mathrm{Id} \otimes \Delta^+\big)\delta^+(\mathcal{I}_n^+ \tau) &= (\mathrm{Id} \otimes \Delta^+ \mathcal{I}_n^+)\delta\tau \\
&= \sum_{\varphi \trianglelefteq \tau} \varphi \otimes \Big((\mathcal{I}_n^+ \otimes \mathrm{Id})\Delta(\tau/^-\varphi) + \sum_{\ell \in \mathbb{N}\times\mathbb{N}^d} \frac{X_+^\ell}{\ell!} \otimes \mathcal{I}_{n+\ell}^+(\tau/^-\varphi)\Big);
\end{aligned}$$

hence, we have

$$\mathcal{M}^{(13)}\big(\delta^+ \otimes \delta^+\big)\Delta^+(\mathcal{I}_n^+ \tau) = \big(\mathrm{Id} \otimes \Delta^+\big)\delta^+(\mathcal{I}_n^+ \tau).$$

It remains to prove (5.4) and (5.6) for elements of the form $X_+^n \tau$. This is elementary using the multiplicative property of S^+ and Δ^+. ■

6. Multi-pre-Lie structure and renormalized equations

Let us summarize the successive steps that we have followed after the formalism of regularity structures was set up in Section 2. We described in Section 3 a particular class of regularity structures, and the class of admissible models on them, that have the property that one can lift the integral operator $(\partial_t - \Delta_x)^{-1}$ into an operator on some spaces of modelled distributions that has a Schauder-type continuity property given in Theorem 17. This result played a crucial role in the local in time well-posedness result proved in Theorem 23 of Section 4. It gives us a modelled distribution $\boldsymbol{u}^{\mathsf{M}}$ that solves a well-defined regularity structure formulation of a fixed-point formulation of an ill-defined singular (stochastic) PDE. We define the (model-dependent) solution of this ill-defined equation as the (model-dependent) reconstruction $u^{\mathsf{M}} = \mathsf{R}^{\mathsf{M}}(\boldsymbol{u}^{\mathsf{M}})$ of $\boldsymbol{u}^{\mathsf{M}}$. To make this definition consistent with the initial objective, we would like to use some models M for which $\Pi(\Xi) = \zeta$. The construction of some admissible model that has this property is made very non-trivial by the fact that ζ has low regularity. As a matter of fact, this cannot be done in a deterministic reasonable way, but one can construct some random admissible models that are limits in a probabilistic sense of some smooth models built from the canonical admissible model M^ε associated with a regularized noise ζ_ε. The construction recipe

$$(\mathsf{M}, k) \mapsto {}^k\mathsf{M}$$

for these renormalized models was given in Theorem 28 in Section 5. Denote by

$$\mathsf{M}^\varepsilon = (\mathrm{g}^\varepsilon, \Pi^\varepsilon)$$

the canonical admissible model associated with a regularized noise ζ_ε. We will see in Section 7 that there is a particular choice of character k_ε for which the ${}^{k_\varepsilon}\Pi^\varepsilon(\tau)(x)$ are centered for all the τ of negative homogeneity and all state space point x. Chandra and Hairer first proved the probabilistic convergence of the renormalized admissible models ${}^{k_\varepsilon}\mathsf{M}^\varepsilon$ to some limit admissible random model $\overline{\mathsf{M}}$. We will not prove this result in the tourist guide and refer the reader to the review [10] for some information on this matter. Rather we will see in the present section that one can give a somewhat explicit description of the dynamics of $u^{\overline{\mathsf{M}}}$ based on the following facts.

(a) The continuity of the map $\mathsf{M} \mapsto u^{\mathsf{M}}$ ensures that the dynamics of $u^{\overline{\mathsf{M}}}$ is the limit of the dynamics of the $\bar{u}_\varepsilon := u^{{}^{k_\varepsilon}\mathsf{M}^\varepsilon}$.

(b) We will see in the present section that $\bar{u}_\varepsilon$ is actually the solution of an explicit stochastic PDE driven by the regularized noise ζ_ε, called the *renormalized equation*. This is the main result of this section, stated in Theorem 43.

We will concentrate in this section on the study of the generalized KPZ equation

$$\begin{aligned}\big(\partial_{x_0} - \Delta_{x'} + 1\big)u &= f(u)\zeta + g_2(u)(\partial_{x'}u)^2 + g_1(u)(\partial_{x'}u) + g_0(u) \\ &= f(u)\zeta + g(u, \partial_{x'}u),\end{aligned} \tag{6.1}$$

with a given initial condition. It already involves the main difficulties of the most general situation, with the advantage of leaving aside a number of purely technical and notational matters compared to the most general situation.

1. Picard iteration and decorated trees. We saw in Section 4 that there is a unique modelled distribution

$$\boldsymbol{u} = \sum_{\tau\in\mathcal{B}} u_\tau \tau \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g}),$$

with $\gamma \in (-\beta_0, 2)$ and $\eta \in (0, \beta_0 + 2]$ solving the lift (4.10) in the regularity structure $\mathcal{T}$ of equation (6.1). It satisfies on the domain $(0, \frac{t_0}{2}) \times \mathbb{T}^d$ the fixed-point problem

$$\begin{aligned}\boldsymbol{u} &\simeq \mathcal{I}\big(f^\star(\boldsymbol{u})\Xi + g_2^\star(\boldsymbol{u}) \star (D\boldsymbol{u})^{\star 2} + g_1^\star(\boldsymbol{u}) \star D\boldsymbol{u} + g_0^\star(\boldsymbol{u})\big) \\ &\simeq \frac{f^{(k)}(u)}{k!} u_{\tau_1}\cdots u_{\tau_k} \mathcal{I}\big(\tau_1\cdots\tau_k\,\Xi\big) + \frac{g_2^{(k)}(u)}{k!} u_{\tau_1}\cdots u_{\tau_k} u_{\sigma_1} u_{\sigma_2} \mathcal{I}\big(\tau_1\cdots\tau_k D_i\sigma_1 D_j\sigma_2\big) \\ &\quad + \frac{g_1^{(k)}(u)}{k!} u_{\tau_1}\cdots u_{\tau_k} u_{\sigma_1} \mathcal{I}\big(\tau_1\cdots\tau_k D_i\sigma_1\big) + \frac{g_0^{(k)}(u)}{k!} u_{\tau_1}\cdots u_{\tau_k} \mathcal{I}\big(\tau_1\cdots\tau_k\big),\end{aligned} \tag{6.2}$$

up to some model-dependent non-trivial polynomial components, with $\tau_k, \sigma_\ell \in \mathcal{B}$, and implicit sums over $\mathcal{B}$ and $i, j \in \{1, \dots, d\}$. We see on this identity that T needs at least to be stable by the operations

$$
\begin{aligned}
&(\tau_1, \dots, \tau_k, \sigma_1, \sigma_2) \\
&\quad \mapsto \mathcal{I}(\tau_1 \cdots \tau_k), \mathcal{I}(\tau_1 \cdots \tau_k \Xi), \mathcal{I}(\tau_1 \cdots \tau_k D_i \sigma_1), \mathcal{I}(\tau_1 \cdots \tau_k D_i \sigma_1 D_j \sigma_2);
\end{aligned}
$$

this naturally endows the elements of T with a tree/inductive structure. This fact is common to all the equations that can be treated by the methods of regularity structures. This leads us in Section 6.1 to setting the framework of *rooted decorated trees* as a convenient encoding of the elements of T.

2. Decorated trees and pre-Lie algebras. The importance of this algebraic setting comes from the fact that the vector space V spanned by the set of all rooted trees with vertex and edge decorations in some given sets happens to be a universal object in a class of algebraic structures called multi-pre-Lie algebras. Morphisms of such multi-pre-Lie algebras defined on V are thus determined by their restrictions to a set of generators. We show in Section 6.2 that the modelled distribution solution of the regularity structure lift of equation (6.1) involves precisely such a morphism, with values in the space of vector fields; see Proposition 36.

3. Pre-Lie algebras and renormalization. The regularity structure associated with equation (6.1) is built from V, with T and T^+ subsets of V. Building $\mathcal{T}$ within V, any renormalization structure $\mathcal{U}$ compatible with $\mathcal{T}$ and satisfying Assumption C will also be built within V, with U and U^- some subsets of V. Theorem 43 below shows that $\bar{u}_\varepsilon$ is the solution of an explicit equation driven by ζ_ε. This result was first proved in the seminal work [17] of Bruned, Chandra, Chevyrev, and Hairer. The proof builds on the fact that the dual renormalization map $\widetilde{k}^*$ that one can associate to any $k \in G^-_{\mathrm{ad}}$ happens to be a multi-pre-Lie morphism under a *compatibility condition* on the multi-pre-Lie structure and the renormalization operator δ, found here under the form of Assumption D3.

Assumptions D1–D3 to be found in this section are all met in the case of a general subcritical system of singular stochastic PDEs, and we verify them by hand in Section 9 where we construct the regularity and renormalization structures associated with the generalized KPZ equation. We emphasize them here as "assumptions" to stress the mechanics at work in the most general case.

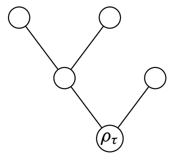
6.1. Free $\mathtt{E}$-multi-pre-Lie algebra generated by $\mathtt{N}$

We introduce in the first paragraph the space of edge and node decorated trees. Decoration spaces are associated to any given system of singular stochastic PDEs, and the associated space of decorated trees provides the background scene from which one can define the regularity and renormalization structures associated with the system. The multi-pre-Lie structure of the space of decorated trees is introduced in another paragraph and its dual operator described explicitly.

Decorated trees.

Definition. Let $\mathfrak{T}_{\mathrm{n}}$ (called a *node-type set*) and $\mathfrak{T}_{\mathrm{e}}$ (called an *edge-type set*) be abstract sets.

- A *rooted tree* τ is a finite connected non-planar graph without loops, with a node set N_τ and an edge set E_τ, and with a distinguished node ρ_τ, called the *root*. The root defines a natural order on each edge, from the root to the leaves. In particular, each edge $e \in E_\tau$ is written as the form $e = (u, v)$, where $u, v \in N_\tau$ are endpoints of e and u is closer to the root. u is called a parent of v, and v is called a child of u.
  We identify two trees τ and σ if they are graph isomorphic, so we always write a graph by putting ancestors lower and descendants upper. The root is put at the bottom. Here is an example.



- A *typed rooted tree* is a rooted tree with type maps $\mathrm{t}_{\mathrm{n}} : N_\tau \to \mathfrak{T}_{\mathrm{n}}$ and $\mathrm{t}_{\mathrm{e}} : E_\tau \to \mathfrak{T}_{\mathrm{e}}$. Moreover, a *rooted decorated tree* is a typed rooted tree τ with two maps

$$\mathfrak{n} : N_\tau \to \mathbb{N}^{d+1}, \quad \mathfrak{e} : E_\tau \to \mathbb{N}^{d+1}.$$

  We denote a generic typed rooted tree by Greek letters like τ, and a generic rooted decorated trees with two decorations $\mathfrak{n}, \mathfrak{e}$ by $\tau_{\mathfrak{e}}^{\mathfrak{n}}$ or a bold letter $\boldsymbol{\tau}$.

We will consider later rooted trees τ equipped with three decorations $\mathfrak{n}$, $\mathfrak{o}$, $\mathfrak{e}$ – see Section 9.1 for the precise definitions. In this section, we hide the $\mathfrak{o}$-decoration in the node-type map, so we consider the type sets

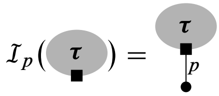
$$\mathfrak{T}_{\mathrm{n}} = \{\bullet, \circ\} \cup \{{\color{red}\bullet}^{\cdot,\alpha}\}_{\alpha \in \mathbb{R}}.$$

The node type $\bullet$ represents the monomial $\mathbf{1} = X^0$, and $\circ$ represents the noise Ξ. The third node type ${\color{red}\bullet}^{\cdot,\alpha}$ is a node with the $\mathfrak{o}$-decoration α. The set $\mathfrak{T}_{\mathrm{e}}$ labels the set of differential operators involved in the system of equations under study. There is a single operator $\partial_{x_0} - \Delta_{x'} + 1$ in the example of the generalized (KPZ) equation (6.1), so the set $\mathfrak{T}_{\mathrm{e}}$ consists of only one element, associated with the integration operator $\mathcal{I}$ in that case. If we consider a system of singular stochastic PDEs involving different operators, different operators $\mathcal{I}$'s would be associated with each of them and the set $\mathfrak{T}_{\mathrm{e}}$ would collect them all.

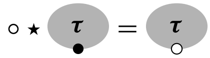
An element $X^n \in T$ is denoted by $\bullet^n$, that is, a graph with only one node with the type $\bullet$ and the $\mathfrak{n}$-decoration $n \in \mathbb{N}^{d+1}$. An edge with $\mathfrak{e}$-decoration $p \in \mathbb{N}^{d+1}$ represents the operator $\mathcal{I}_p$, with the notations of Section 3.2, for one of the operators $\mathcal{I}$ involved in the equation.

All operations appearing in the equation (6.2) are graphically defined as follows. In the following pictures, types and decorations are omitted unless necessary, and the root of $\boldsymbol{\tau}$ in the first bullet and of $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_m$ in the third bullet is denoted by a square.

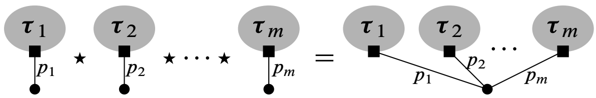
- The integration $\boldsymbol{\tau} \mapsto \mathcal{I}_p(\boldsymbol{\tau})$ is given by the map connecting the root of $\boldsymbol{\tau}$ with a new node, which becomes a root of the tree $\mathcal{I}_p(\boldsymbol{\tau})$, and giving the $\mathfrak{e}$-decoration $p \in \mathbb{N}^{d+1}$ to the connecting edge.



- The product $\mathfrak{b}^n \star \boldsymbol{\tau}$ for $\mathfrak{b} \in \mathfrak{T}_{\mathrm{n}}$, $n \in \mathbb{N}^{d+1}$, and $\boldsymbol{\tau}$ with $\mathfrak{t}_{\mathrm{n}}(\rho_\tau) = \bullet$ and $\mathfrak{n}(\rho_\tau) = 0$ is given changing the node type of ρ_τ to $\mathfrak{b}$ and $\mathfrak{n}$-decoration to n. For example, if $\mathfrak{b} = \circ$ and $n = 0$,



- The product of trees $\mathcal{I}_{p_j}(\boldsymbol{\tau}_j)$ $(j = 1, \dots, m)$ is given by the *tree product*, that is joining their roots.



Thus, we see that the rooted trees obtained by the above operations are sufficient to describe the fixed-point problem (6.2). The symbol $\bullet^{\cdot,\alpha}$ does not come from the fixed-point problem (6.2), but its use is made clear in Section 6.3. As we concentrate in this section on the generalized (KPZ) equation, the edge-type set $\mathfrak{T}_{\mathrm{e}}$ will consist of a single element, suggestively denoted by $\mathcal{I}$. There is no difficulty in working with a finite edge-type set.

Definition. Let $\mathcal{V}$ be the set of *all* rooted decorated trees with type sets $\mathfrak{T}_{\mathrm{n}} = \{\bullet, \circ\} \cup \{\bullet^{\cdot,\alpha}\}_{\alpha \in \mathbb{R}}$ and $\mathfrak{T}_{\mathrm{e}} = \{\mathcal{I}\}$, and let V be the vector space spanned by $\mathcal{V}$. Moreover, denote by $(\boldsymbol{\tau}^* : V \to \mathbb{R})_{\boldsymbol{\tau} \in \mathcal{V}}$ the dual basis of $\mathcal{V}$, and let V^* be the vector space spanned by $\{\boldsymbol{\tau}^*\}_{\boldsymbol{\tau} \in \mathcal{V}}$.

Throughout this section, we view each element of $\mathcal{V}$ as the rooted tree τ with the composite decorations $(\mathfrak{t}_{\mathrm{n}}, \mathfrak{n}) : N_\tau \to \mathsf{N}$ and $(\mathfrak{t}_{\mathrm{e}}, \mathfrak{e}) : E_\tau \to \mathsf{E}$, where

$$\mathsf{E} := \mathfrak{T}_{\mathrm{e}} \times \mathbb{N}^{d+1} \simeq \mathbb{N}^{d+1}, \quad \mathsf{N} := \mathfrak{T}_{\mathrm{n}} \times \mathbb{N}^{d+1}.$$

The set N is considered as a subset of $\mathcal{V}$ consisting of simple trees

$$\mathsf{N} = \{\mathfrak{b}^n\}_{\mathfrak{b} \in \mathfrak{T}_{\mathrm{n}}, n \in \mathbb{N}^{d+1}} = \left\{\circ^\ell, \bullet^m, \bullet^{n,\alpha}\right\}_{\ell,m,n \in \mathbb{N}^{d+1}, \alpha \in \mathbb{R}}.$$

Write $\mathsf{N}^0 := \{\mathfrak{b}^0\}_{\mathfrak{b} \in \mathfrak{T}_{\mathrm{n}}} \simeq \mathfrak{T}_{\mathrm{n}}$. We introduce a few notations. Note that each $\boldsymbol{\tau} \in \mathcal{V}$ has a decomposition of the form

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i) = \mathfrak{b}^n \star \mathcal{I}_{p_1}(\boldsymbol{\tau}_1) \star \cdots \star \mathcal{I}_{p_a}(\boldsymbol{\tau}_a)$$

with $\mathfrak{b}^n \in \mathsf{N}$, $p_1, \dots, p_a \in \mathbb{N}^{d+1}$, and $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_a \in \mathcal{V}$. Taking care of the number of automorphisms of $\boldsymbol{\tau}$ that leave it fixed, for $\boldsymbol{\tau}$ of the form

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \mathop{\bigstar}_{j=1}^{b} \left(\mathcal{I}_{q_j}(\boldsymbol{\sigma}_j)\right)^{\star m_j},$$

with $(q_i, \boldsymbol{\sigma}_i) \neq (q_j, \boldsymbol{\sigma}_j)$ for any $i \neq j$, define inductively

$$S(\boldsymbol{\tau}) := n! \prod_{j=1}^{b} S(\boldsymbol{\sigma}_j)^{m_j}\, m_j!.$$

Then, we define the paring $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ between V and V^* by

$$\langle\!\langle \boldsymbol{\tau}, \boldsymbol{\sigma}^* \rangle\!\rangle := S(\boldsymbol{\sigma})\, \boldsymbol{\sigma}^*(\boldsymbol{\tau}) \tag{6.3}$$

for $\boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathcal{V}$. We see V^* as a part of the algebraic dual of V. (As V is infinite dimensional, V^* is not equal to the full algebraic dual of V.) The “copy” space V^* will play an important role in the second half part of this section.

Canonical model. Given a *smooth* noise $\zeta \in \mathcal{C}^\infty(\mathbb{R} \times \mathbb{R}^d)$, we define the canonical operator Π^ζ on the whole of V requiring that it is multiplicative with respect to the $\star$ product and setting, for all $x \in \mathbb{R} \times \mathbb{R}^d$,

$$\Pi^\zeta(\circ^n)(x) = x^n \zeta(x), \quad \Pi^\zeta(\bullet^n)(x) = \Pi^\zeta({\color{red}\bullet}^{n,\alpha})(x) = x^n,$$

and

$$\Pi^\zeta(\mathcal{I}_p \boldsymbol{\tau}) = \partial^p \mathbf{K}(\Pi^\zeta \boldsymbol{\tau}),$$

for all $n, \alpha, \boldsymbol{\tau}, p$. The regularity structures we will work with have spaces T and T^+ that are subsets of V. Since all functions $\Pi^\zeta \boldsymbol{\tau}$ are smooth, the restriction of Π^ζ to $T_{<0}$ defines the *canonical model*

$$\mathsf{M}^\zeta = \big(\mathsf{g}^\zeta, \Pi^\zeta\big)$$

on the regularity structure $\mathcal{T}$, from Theorem 19. Things are explicit here as the multiplicativity and the $\mathbf{K}$-admissibility properties fix the definition of Π^ζ on all decorated trees in V. Emphasize the fact that since the map Π^ζ is multiplicative, its associated reconstruction map is also multiplicative.

Multi-pre-Lie algebras. We first recall the definition of a multi-pre-Lie algebra, referring the reader to Foissy’s article [41] for basics on multi-pre-Lie algebras. All we need to know on the subject is the following definition and the result of Proposition 32 below.

Definition. Let E be a set. A vector space W, equipped with a family $(\rhd_{\mathsf{e}})_{\mathsf{e} \in \mathsf{E}}$ of bilinear maps from $W \times W$ into W, is called an E-*multi-pre-Lie algebra* if one has

$$(a \rhd_{\mathsf{e}} b) \rhd_{\mathsf{e}'} c - a \rhd_{\mathsf{e}} (b \rhd_{\mathsf{e}'} c) = (b \rhd_{\mathsf{e}'} a) \rhd_{\mathsf{e}} c - b \rhd_{\mathsf{e}'} (a \rhd_{\mathsf{e}} c)$$

for all $a, b, c \in W$, and $\mathsf{e}, \mathsf{e}' \in \mathsf{E}$.

The two arguments of a pre-Lie product $a \rhd_{\mathsf{e}} b$ do not play a symmetric role, and we think here of a as acting on b via the operator $\rhd_{\mathsf{e}}$; we read $a \rhd_{\mathsf{e}} b$ from left to right. Here is an example of E-multi-pre-Lie algebra. Take E finite, identified with $\{1, \dots, |\mathsf{E}|\}$, and consider the space of smooth functions on $\mathbb{R}^{|\mathsf{E}|}$. Then, the family of differentiation operators

$$G \triangleright_{\mathsf{e}} H := G \partial_{x_{\mathsf{e}}} H$$

defines an E-multi-pre-Lie algebra. If E consists of a single element $\triangleright$, this operator is called a *pre-Lie product*, and a vector space equipped with a pre-Lie product is called a *pre-Lie algebra*. Any pre-Lie algebra is Lie-admissible, in the sense that the map $(a,b) \mapsto a \triangleright b - b \triangleright a$ defines a Lie bracket. The relevance of the multi-pre-Lie structure in the study of singular stochastic PDEs comes from Proposition 36 in the next section, as it identifies the components u_τ of solutions $\boldsymbol{u} = \sum u_\tau \tau$ regularity structures lifts of a singular stochastic PDEs as E-multi-pre-Lie algebra morphisms.

We define now the multi-pre-Lie structure in the space V^*. The reason for working on V^* rather than on V will appear clearly in Sections 6.2 and 6.3. With the spaces V and V^* being infinite dimensional, the symbol $\otimes$ denotes below the algebraic tensor product of these spaces with themselves, without any completion.

Definition. Given $\mathsf{e} \in \mathsf{E} \simeq \mathbb{N}^{d+1}$, a node v of a decorated tree $\boldsymbol{\sigma} \in \mathcal{V}$ and $\boldsymbol{\tau} \in \mathcal{V}$, denote by

$$\boldsymbol{\tau} \overset{\mathsf{e}}{\rightarrow}_{(v)} \boldsymbol{\sigma}$$

the element of $\mathcal{V}$ obtained by grafting $\boldsymbol{\tau}$ on the node v of $\boldsymbol{\sigma}$, along an edge of e-decoration e. Define also

$$\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_{(v)} \sigma_{\mathfrak{e}}^{\mathfrak{n}} := \sum_{m \in \mathbb{N}^{d+1}; m \leq \mathfrak{n}(v) \wedge p_{\mathsf{e}}} \binom{\mathfrak{n}(v)}{m} \boldsymbol{\tau} \xrightarrow{\mathsf{e}-m}_{(v)} \sigma_{\mathfrak{e}}^{\mathfrak{n}-m\mathbf{1}_v} \in V,$$

$$\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \sigma_{\mathfrak{e}}^{\mathfrak{n}} := \sum_{v \in N_\sigma} \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_{(v)} \sigma_{\mathfrak{e}}^{\mathfrak{n}} \in V,$$

where $\mathbf{1}_v$ is the indicator function of v. Recall that the binomial coefficient of multi-indices is defined at the end of Section 1. Finally, define a linear map

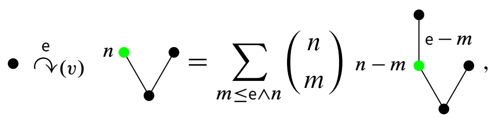
$$\overset{\mathsf{e}}{\curvearrowright} : V^* \otimes V^* \to V^*$$

by

$$\boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}^* := (\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma})^*, \qquad \boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathcal{V},$$

where the map $(\cdot)^* : V \to V^*$ is the linear extension of the map $\mathcal{V} \ni \boldsymbol{\tau} \mapsto \boldsymbol{\tau}^* \in V^*$.

Here is an example:



where v is colored in green. The next statement is fundamental and can be proved as Corollary 9 in Foissy's work [41] – it was first proved in Proposition 4.21 of Bruned, Chandra, Chevyrev, and Hairer's work [17]. A proof can be found in Appendix C.2.

Proposition 31. *The space V^* with the operators $\{\overset{\mathsf{e}}{\curvearrowright}\}_{\mathsf{e}\in\mathsf{E}}$ is the free* E*-multi-pre-Lie algebra generated by* N*, in the sense that the universal property (b) in Appendix C.2 holds.*

Any morphism from V^* into an E-multi-pre-Lie algebra is thus determined by its restriction to the generators N of V^*. This is the universal property of the free E-multi-pre-Lie algebra with generators N. In particular, if two E-multi-pre-Lie morphisms from V^* into the same E-multi-pre-Lie algebra coincide on the generators of V^*, then they are equal.

The space $T = U$ of the regularity and renormalization structures associated with the generalized KPZ equation is a subspace of V with each space T_β, T_α^+, $U_{\beta'}$, $U_{\alpha'}^-$ spanned by finitely many rooted decorated trees. Denote by $\pi_U : V \to U$ the canonical projection. The next assumption is a piece of properties to be satisfied by the basis $\mathcal{B}$ of T and U. In Section 9.1, $\mathcal{B}$ is defined as the set of all trees *strongly conforming* to the rule. The first one means that the $\mathfrak{n}$-decoration is independent of the rule and the second one means that the rule is local. The last one describes that the projection map π_U behaves consistently with respect to all the grafting products $\overset{\mathsf{e}}{\curvearrowright}$.

Assumption D1. The homogeneous basis $\mathcal{B}$ of T and U is a subset of $\mathcal{V}$ with the following properties. (Recall that the notions of homogeneity in T and U are different.)

- If $\boldsymbol{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathcal{B}$, then $\tau_{\mathfrak{e}}^{\mathfrak{m}} \in \mathcal{B}$ for any $\mathfrak{m} : N_\tau \to \mathbb{N}^{d+1}$.
- If $\boldsymbol{\tau} = \mathfrak{b}^n \star \bigstar_{i=1}^a \mathcal{I}_{p_i}(\boldsymbol{\tau}_i) \in \mathcal{B}$, then $\mathfrak{b}, \boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_a \in \mathcal{B}$.
- For any $\boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathcal{V}$ and $\mathsf{e} \in \mathsf{E}$,

$$\pi_U\big(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} (\pi_U \boldsymbol{\sigma})\big) = \pi_U\big((\pi_U \boldsymbol{\tau}) \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}\big) = \pi_U\big(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}\big).$$

Set

$$U^* := \operatorname{span}\{\boldsymbol{\tau}^*; \boldsymbol{\tau} \in \mathcal{B}\}$$

and denote by $\pi_{U^*} : V^* \to U^*$ the canonical projection. Then, we define the map

$$\overset{\mathsf{e}}{\curvearrowright}_{\flat} : U^* \otimes U^* \to U^*,$$

setting

$$\boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright}_{\flat} \boldsymbol{\sigma}^* := \pi_{U^*}\big(\boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}^*\big).$$

The following statement is proved in Appendix C.2.

Proposition 32. *Under Assumption D1, the space U^* with the operators $\{\overset{\mathsf{e}}{\curvearrowright}_{\flat}\}_{\mathsf{e}\in\mathsf{E}_{\leq 1}}$ is the* E*-multi-pre-Lie algebra generated by* $\mathsf{N} \cap \mathcal{B}$.

Finally, we define an operator playing the role of "antiderivative".

Definition. For each $i \in \{0, 1, \dots, d\}$, define the linear map $\uparrow_i : V^* \to V^*$ by

$$\uparrow_i (\tau_{\mathfrak{e}}^{\mathfrak{n}})^* = \sum_{v \in N_\tau} (\tau_{\mathfrak{e}}^{\mathfrak{n}+e_i \mathbf{1}_v})^*.$$

The map $\uparrow_i$ sends U^* into itself under Assumption D1. We denote by

$$\downarrow_i \colon V \to V$$

the dual map of $\uparrow_i \colon V^* \to V^*$ under the pairing (6.3), that is,

$$\langle\!\langle \downarrow_i \boldsymbol{\tau}, \boldsymbol{\sigma}^* \rangle\!\rangle = \langle\!\langle \boldsymbol{\tau}, \uparrow_i \boldsymbol{\sigma}^* \rangle\!\rangle$$

for any $\boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathcal{V}$. Moreover, we extend the pairing (6.3) into a pairing between $V \otimes V$ and $V^* \otimes V^*$, setting

$$\langle\!\langle \boldsymbol{\tau}_1 \otimes \boldsymbol{\tau}_2, \boldsymbol{\sigma}_1^* \otimes \boldsymbol{\sigma}_2^* \rangle\!\rangle := \langle\!\langle \boldsymbol{\tau}_1, \boldsymbol{\sigma}_1^* \rangle\!\rangle \, \langle\!\langle \boldsymbol{\tau}_2, \boldsymbol{\sigma}_2^* \rangle\!\rangle.$$

Under such pairings, denote by

$$\upharpoonleft\!\downharpoonright_{\mathsf{e}} \colon V \to V \otimes V$$

the dual map of $\overset{\mathsf{e}}{\curvearrowright} \colon V^* \otimes V^* \to V^*$, that is,

$$\langle\!\langle \upharpoonleft\!\downharpoonright_{\mathsf{e}} \boldsymbol{\eta}, \boldsymbol{\tau}^* \otimes \boldsymbol{\sigma}^* \rangle\!\rangle := \langle\!\langle \boldsymbol{\eta}, \boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}^* \rangle\!\rangle \tag{6.4}$$

for any $\boldsymbol{\tau}, \boldsymbol{\sigma}, \boldsymbol{\eta} \in \mathcal{V}$ and $\mathsf{e} \in \mathsf{E}$. The following explicit formulas for $\downarrow_i$ and $\upharpoonleft\!\downharpoonright_{\mathsf{e}}$ are helpful to get a graphical image. It is used only in the proof of Theorem 47 giving an explicit construction of the regularity and renormalization structures associated with the generalized KPZ equation.

Lemma 33. *For any $i \in \{0, 1, \dots, d\}$ and any $\boldsymbol{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathcal{V}$, one has*

$$\downarrow_i \left(\tau_{\mathfrak{e}}^{\mathfrak{n}}\right) = \sum_{v \in N_\tau,\, e_i \le \mathfrak{n}(v)} \mathfrak{n}(v)\, \tau_{\mathfrak{e}}^{\mathfrak{n} - e_i \mathbf{1}_v}.$$

Moreover, for any $\boldsymbol{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}} \in \mathcal{V}$ and any $\mathsf{e} \in E$, one has

$$\upharpoonleft\!\downharpoonright_{\mathsf{e}} (\tau_{\mathfrak{e}}^{\mathfrak{n}}) = \sum_{e=(v,w) \in E_\tau; \mathfrak{e}(e) \le \mathsf{e}} \frac{1}{(\mathsf{e} - \mathfrak{e}(e))!} (C_e \tau)_{\mathfrak{e}}^{\mathfrak{n}} \otimes (P_e \tau)_{\mathfrak{e}}^{\mathfrak{n} + (\mathsf{e} - \mathfrak{e}(e)) \mathbf{1}_v}, \tag{6.5}$$

where $C_e \tau$ and $P_e \tau$ are the two connected components of the graph $\tau \setminus \{e\}$, with $P_e \tau$ containing the root of τ. (Again, recall that the factorial of a multi-index is defined at the end of Section 1.)

Proof. We show that equation (6.4) holds for the map $\upharpoonleft\!\downharpoonright_{\mathsf{e}}$ defined by the second formula (6.5). The first formula is proved by a similar argument. Note that, for any elements $\boldsymbol{\tau} = \mathfrak{b}^n \star \bigstar_{i=1}^a \mathcal{I}_{p_i}(\boldsymbol{\tau}_i) \in \mathcal{V}$ and $\boldsymbol{\sigma} = \mathfrak{u}^m \star \bigstar_{j=1}^b \mathcal{I}_{q_j}(\boldsymbol{\sigma}_j) \in \mathcal{V}$, one has

$$\langle\!\langle \boldsymbol{\tau}, \boldsymbol{\sigma}^* \rangle\!\rangle = \mathbf{1}_{\mathfrak{b} = \mathfrak{u}, n = m, a = b}\, n! \sum_{s \in S_a} \prod_{i=1}^a \mathbf{1}_{p_i = q_{s(i)}} \langle\!\langle \boldsymbol{\tau}_i, \boldsymbol{\sigma}_{s(i)}^* \rangle\!\rangle, \tag{6.6}$$

where S_a is the symmetric group of the set $\{1, 2, \dots, a\}$. For $\boldsymbol{\eta}$ of the form

$$\boldsymbol{\eta} = \mathfrak{b}^n \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\eta}_i),$$

we divide the right-hand side of (6.5) according to whether the edge e is connected to the root or not and have

$$\begin{aligned} \upharpoonleft\!\downharpoonright_{\mathsf{e}} \boldsymbol{\eta} &= \sum_i \frac{1}{(\mathsf{e} - p_i)!} \boldsymbol{\eta}_i \otimes \mathfrak{b}^{n+\mathsf{e}-p_i} \star \mathop{\bigstar}_{j:j\neq i} \mathcal{I}_{p_j}(\boldsymbol{\eta}_j) \\ &\quad + \sum_i \sum_{(\boldsymbol{\eta}_i)} \boldsymbol{\eta}_i^1 \otimes \mathfrak{b}^n \star \mathcal{I}_{p_i}(\boldsymbol{\eta}_i^2) \star \mathop{\bigstar}_{j:j\neq i} \mathcal{I}_{p_j}(\boldsymbol{\eta}_j) \\ &=: \upharpoonleft\!\downharpoonright_{\mathsf{e}}^1 \boldsymbol{\eta} + \upharpoonleft\!\downharpoonright_{\mathsf{e}}^2 \boldsymbol{\eta}, \end{aligned}$$

where we write $\upharpoonleft\!\downharpoonright_{\mathsf{e}} \boldsymbol{\eta}_i = \sum_{(\boldsymbol{\eta}_i)} \boldsymbol{\eta}_i^1 \otimes \boldsymbol{\eta}_i^2$ following Sweedler's notation in the first equality. Similarly, for any $\boldsymbol{\tau} \in \mathcal{V}$ and $\boldsymbol{\sigma} = \mathfrak{u}^m \star \bigstar_{j=1}^{b} \mathcal{I}_{q_j}(\boldsymbol{\sigma}_j) \in \mathcal{V}$, one has

$$\begin{aligned} \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma} &= \sum_\ell \binom{m}{\ell} \mathfrak{u}^{m-\ell} \star \mathcal{I}_{\mathsf{e}-\ell}(\boldsymbol{\tau}) \star \mathop{\bigstar}_{j=1}^{b} \mathcal{I}_{q_j}(\boldsymbol{\sigma}_j) \\ &\quad + \sum_{j=1}^{b} \mathfrak{u}^m \star \mathcal{I}_{q_j}\big(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}_j\big) \star \mathop{\bigstar}_{k;k\neq j} \mathcal{I}_{q_k}(\boldsymbol{\sigma}_k) \\ &=: \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_1 \boldsymbol{\sigma} + \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_2 \boldsymbol{\sigma}. \end{aligned}$$

Hence, it is sufficient to show that

$$\langle\!\langle \boldsymbol{\eta}, \boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright}_1 \boldsymbol{\sigma}^* \rangle\!\rangle = \langle\!\langle \upharpoonleft\!\downharpoonright_{\mathsf{e}}^1 \boldsymbol{\eta}, \boldsymbol{\tau}^* \otimes \boldsymbol{\sigma}^* \rangle\!\rangle, \tag{6.7}$$

$$\langle\!\langle \boldsymbol{\eta}, \boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright}_2 \boldsymbol{\sigma}^* \rangle\!\rangle = \langle\!\langle \upharpoonleft\!\downharpoonright_{\mathsf{e}}^2 \boldsymbol{\eta}, \boldsymbol{\tau}^* \otimes \boldsymbol{\sigma}^* \rangle\!\rangle. \tag{6.8}$$

It is not difficult to show (6.7) directly from (6.6). For (6.8), it is sufficient to consider $\boldsymbol{\sigma} = \mathfrak{b}^n \star \bigstar_{i=1}^{a} \mathcal{I}_{q_i}(\boldsymbol{\sigma}_i)$, and for such $\boldsymbol{\sigma}$ one has

$$\langle\!\langle \boldsymbol{\eta}, \boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright}_2 \boldsymbol{\sigma}^* \rangle\!\rangle = n! \sum_{i=1}^{a} \sum_{s\in S_a} \mathbf{1}_{q_{s(i)}=p_i} \langle\!\langle \boldsymbol{\eta}_i, \boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}_{s(i)}^* \rangle\!\rangle \prod_{j;j\neq i} \mathbf{1}_{q_{s(j)}=p_j} \langle\!\langle \boldsymbol{\eta}_j, \boldsymbol{\sigma}_{s(j)}^* \rangle\!\rangle$$

and

$$\begin{aligned} &\langle\!\langle \upharpoonleft\!\downharpoonright_{\mathsf{e}}^2 \boldsymbol{\eta}, \boldsymbol{\tau}^* \otimes \boldsymbol{\sigma}^* \rangle\!\rangle \\ &\quad = n! \sum_{i=1}^{a} \sum_{(\boldsymbol{\eta}_i)} \sum_{s\in S_a} \langle\!\langle \boldsymbol{\eta}_i^1, \boldsymbol{\tau}^* \rangle\!\rangle \mathbf{1}_{q_{s(i)}=p_i} \langle\!\langle \boldsymbol{\eta}_i^2, \boldsymbol{\sigma}_{s(i)}^* \rangle\!\rangle \prod_{j;j\neq i} \mathbf{1}_{q_{s(j)}=p_j} \langle\!\langle \boldsymbol{\eta}_j, \boldsymbol{\sigma}_{s(j)}^* \rangle\!\rangle. \end{aligned}$$

Since

$$\sum_{(\boldsymbol{\eta}_i)} \langle\!\langle \boldsymbol{\eta}_i^1, \boldsymbol{\tau}^* \rangle\!\rangle \, \langle\!\langle \boldsymbol{\eta}_i^2, \boldsymbol{\sigma}_{s(i)}^* \rangle\!\rangle = \langle\!\langle 1\!\!\downharpoonright_{\mathsf{e}} \boldsymbol{\eta}, \boldsymbol{\tau}^* \otimes \boldsymbol{\sigma}_{s(i)}^* \rangle\!\rangle,$$

identity (6.8) follows if (6.4) holds for $\boldsymbol{\sigma} = \boldsymbol{\sigma}_i$, which leads to an induction on the number of edges contained in $\boldsymbol{\sigma}$. The case $\boldsymbol{\sigma} = \mathfrak{b}^n \in \mathsf{N}$ is an easy exercise. ■

6.2. Modelled distributions solutions of singular PDEs

The approximate description (6.2) of the fixed-point problem (4.11) leads to an explicit formula for the coefficients of the solution $\boldsymbol{u}$. Noting that $\gamma \in (-\beta_0, 2)$ can be arbitrarily chosen, the solution $\boldsymbol{u}$ of (6.2) is of the form

$$\boldsymbol{u} = \sum_{|k|_{\mathfrak{s}} < \gamma} \frac{u_k}{k!} X^k + \sum_{\boldsymbol{\tau} \in \mathcal{B}, |\boldsymbol{\tau}| < \gamma - 2} u_{\mathcal{I}(\boldsymbol{\tau})} \, \mathcal{I}(\boldsymbol{\tau}). \tag{6.9}$$

Inserting such an expansion into (6.2), we see that all coefficients $u_{\mathcal{I}(\boldsymbol{\tau})}$ are cylindrical functions of

$$\mathsf{u} := (u_k)_{k \in \mathbb{N}^{1+d}} \in \mathbb{R}^{\mathbb{N}^{1+d}}.$$

Here, we say that a function of u is cylindrical if it depends only on a finite number of entries among $(u_k)_{k \in \mathbb{N}^{1+d}}$. For any smooth cylindrical function F, we denote by $\partial_k f := \frac{\partial}{\partial u_k} F$ the derivative with respect to u_k. Moreover, we define the derivative operators $(D_i)_{i=0}^d$ by setting

$$D_i F := \sum_{k \in \mathbb{N}^{1+d}} u_{k+e_i} \partial_k F,$$

and $D^n := \prod_{i=0}^d D_i^{n_i}$ for $n = (n_i)_{i=0}^d \in \mathbb{N}^{d+1}$.

Definition. Set $\mathsf{u}_0 = u_0$ and $\mathsf{u}_1 = (u_{e_i})_{i=1}^d$. Define the linear map F from V^* to the space of cylindrical functions of u as follows. For the primitive trees in $\mathsf{N}^0 := \{\mathfrak{b}^0\}_{\mathfrak{b} \in \mathfrak{T}_{\mathrm{n}}} \simeq \mathfrak{T}_{\mathrm{n}}$, set

$$\begin{aligned} F(\circ^*)(\mathsf{u}) &:= f(\mathsf{u}_0), \\ F(\bullet^*)(\mathsf{u}) &:= g(\mathsf{u}_0, \mathsf{u}_1) := g_2(\mathsf{u}_0)(\mathsf{u}_1)^2 + g_1(\mathsf{u}_0)\mathsf{u}_1 + g_0(\mathsf{u}_0) \\ &:= \sum_{i,j=1}^d g_2^{ij}(u_0) u_{e_i} u_{e_j} + \sum_{i=1}^d g_1^i(u_0) u_{e_i} + g_0(u_0), \\ F\big((\bullet^{0,\alpha})^*\big)(\mathsf{u}) &:= 0. \end{aligned}$$

For a generic tree

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \bigstar_{i=1}^a \mathcal{I}_{p_i}(\boldsymbol{\tau}_i),$$

define inductively

$$F(\boldsymbol{\tau}^*)(\mathsf{u}) := \left\{ \prod_{i=1}^a F(\boldsymbol{\tau}_i^*)(\mathsf{u}) \right\} \left\{ D^n \prod_{i=1}^a \partial_{p_i} \right\} F\big(\mathfrak{b}^*\big)(\mathsf{u}). \tag{6.10}$$

Here are some examples. Recall that • represents $\mathbf{1} = X^0$, and ∘ represents the noise Ξ:

$$F\big((\Xi \star \mathcal{I}(\Xi))^*\big)(\mathsf{u}) = f(u_0)f'(u_0),$$

$$F\big((\mathcal{I}_{e_i}(\Xi) \star \mathcal{I}_{e_j}(\Xi))^*\big)(\mathsf{u}) = \begin{cases} g_2^{ij}(u_0)f(u_0)^2 & (i \neq j), \\ 2g_2^{ii}(u_0)f(u_0)^2 & (i = j). \end{cases}$$

We recall some useful formulas for the derivative operators.

Lemma 34. *For any smooth cylindrical function F of* u *and any $n \in \mathbb{N}^{1+d}$, one has*

$$\frac{D^n F}{n!} = \sum_{\substack{m:\mathbb{N}^{1+d}\times(\mathbb{N}^{1+d}\setminus\{0\})\to\mathbb{N} \\ \sum_q(\sum_k m(k,q))q=n}} \left\{ \prod_{\substack{k\in\mathbb{N}^{1+d} \\ q\in\mathbb{N}^{1+d}\setminus\{0\}}} \frac{1}{m(k,q)!}\left(\frac{u_{k+q}}{q!}\right)^{m(k,q)} \right\} \times \Bigg(\prod_{\substack{k\in\mathbb{N}^{1+d} \\ q\in\mathbb{N}^{1+d}\setminus\{0\}}} \partial_k^{m(k,q)} \Bigg) F \tag{6.11}$$

(Faà di Bruno formula from [17, Lemma A.1]*). Here,* $0! = 1$, $u_k^0 = 1$, *and* $\partial_k^0 = \mathrm{Id}$ *by convention, so the sum and the multiplications on the right-hand side are over only finitely many parameters. Moreover, for any $k, n \in \mathbb{N}^{1+d}$, one has*

$$\partial_k D^n F = \sum_{\ell\in\mathbb{N}^{1+d}} \binom{n}{\ell} D^{n-\ell}\partial_{k-\ell}F, \tag{6.12}$$

where $\binom{n}{\ell} = 0$ if $\ell > n$ and $\partial_{k-\ell} = 0$ if $\ell > k$ by convention, so the sums are over $\ell \leq n \wedge k$.

Proof. The proof of (6.11) is an induction on n. In the case $n = e_i$, the right-hand side of (6.11) coincides with $D_i F$ by definition. Next, assume that (6.11) holds for n and consider $n + e_i$. For simplicity, we denote by M_n the set of all maps

$$m : \mathbb{N}^{1+d} \times (\mathbb{N}^{1+d} \setminus \{0\}) \to \mathbb{N}$$

such that $\sum_q(\sum_k m(k,q))q = n$, and write

$$A(m) = \prod_{k,q} \frac{1}{m(k,q)!}\left(\frac{u_{k+q}}{q!}\right)^{m(k,q)} \quad \text{and} \quad B(m) = \Big(\prod_{k,q} \partial_k^{m(k,q)}\Big)F.$$

By Leibniz rule, we can divide $D_i \frac{D^n F}{n!}$ into two terms according to the fact that D_i is applied to $A(m)$ or $B(m)$. By definition of $D_i = \sum_{\ell\in\mathbb{N}^{1+d}} u_{\ell+e_i}\partial_\ell$, the latter part is reorganized as

$$\sum_{m\in M_n} \sum_{\ell\in\mathbb{N}^{1+d}} \big(m(\ell, e_i) + 1\big)A(\tilde{m}_\ell)B(\tilde{m}_\ell), \tag{6.13}$$

where $\tilde{m}_\ell \in M_{n+e_i}$ is defined by $\tilde{m}_\ell(k,q) = m(k,q) + \mathbf{1}_{(k,q)=(\ell,e_i)}$. On the other hand, since $D_i u_k^m = m u_k^{m-1} u_{k+e_i}$, the former part is also reorganized as

$$\sum_{m\in M_n} \sum_{\ell\in\mathbb{N}^{1+d},\, p\in\mathbb{N}^{1+d}\setminus\{0\}; m(\ell,p)\geq 1} (p_i+1)\big(m(\ell,p+e_i)+1\big) A(\tilde{m}_{\ell,p}) B(\tilde{m}_{\ell,p}), \quad (6.14)$$

where $\tilde{m}_{\ell,p} \in M_{n+e_i}$ is defined by $\tilde{m}_{\ell,p}(k,q) = m(k,q) - \mathbf{1}_{(k,q)=(\ell,p)} + \mathbf{1}_{(k,q)=(\ell,p+e_i)}$. The term (6.13) can be absorbed into the sum (6.14) where the condition on p is replaced by "$p \in \mathbb{N}^{1+d}$", by setting $\tilde{m}_{k,0} := \tilde{m}_k$ and $m(\ell,0) = 1$ for any ℓ. Conversely, for any $\mu \in M_{n+e_i}$, if $\mu(\ell,p) \geq 1$ for some $\ell \in \mathbb{N}^{1+d}$ and $p \in \mathbb{N}^{1+d} \setminus \{0\}$, then there exists a unique $m \in M_n$ such that $\tilde{m}_{\ell,p-e_i} = \mu$. Therefore, the sum (6.14) is reorganized as

$$\begin{aligned} D_i \frac{D^n F}{n!} &= \sum_{m\in M_n} \sum_{\ell,p\in\mathbb{N}^{1+d}; m(\ell,p)\geq 1} (p_i+1)\big(m(\ell,p+e_i)+1\big) A(\tilde{m}_{\ell,p}) B(\tilde{m}_{\ell,p}) \\ &= \sum_{\mu\in M_{n+e_i}} \Bigg(\sum_{m\in M_n,\, \ell,p\in\mathbb{N}^{1+d}; \tilde{m}_{\ell,p}=\mu} (p_i+1)\big(m(\ell,p+e_i)+1\big)\Bigg) A(\mu) B(\mu). \end{aligned}$$

It turns out that the quantity inside the large parentheses is equal to

$$\sum_{\ell\in\mathbb{N}^{1+d}, q\in\mathbb{N}^{1+d}\setminus\{0\}} q_i \mu(\ell,q) = n_i + 1$$

because of the condition that $\mu \in M_{n+e_i}$. Thus, we have $\frac{D_i}{n_i+1}\frac{D^n F}{n!} = \sum_{\mu\in M_{n+e_i}} A(\mu) B(\mu)$. This yields that (6.11) holds for any n.

The proof of (6.12) is also an induction on n. The case $n = e_i$ follows from Leibniz rule:

$$\begin{aligned} \partial_k D_i F &= \sum_{\ell\in\mathbb{N}^{1+d}} \partial_k (u_{\ell+e_i} \partial_\ell F) = \sum_{\ell\in\mathbb{N}^{1+d}} (\partial_k u_{\ell+e_i}) \partial_\ell F + \sum_{\ell\in\mathbb{N}^{1+d}} u_{\ell+e_i} \partial_k \partial_\ell F \\ &= \mathbf{1}_{k\geq e_i} \partial_{k-e_i} F + D_i \partial_k F. \end{aligned}$$

Assuming that (6.12) holds for n, we have for $n + e_i$

$$\begin{aligned} \partial_k D^{n+e_i} F &= \partial_k D^n (D_i F) = \sum_\ell \binom{n}{\ell} D^{n-\ell} \partial_{k-\ell} D_i F \\ &= \sum_\ell \binom{n}{\ell} D^{n-\ell} (\mathbf{1}_{k-\ell\geq e_i} \partial_{k-\ell-e_i} F + D_i \partial_{k-\ell} F) \\ &= \sum_{m\in\mathbb{N}^{1+d}} \binom{n}{m-e_i} D^{n+e_i-m} \partial_{k-m} F + \sum_{m\in\mathbb{N}^{1+d}} \binom{n}{m} D^{n+e_i-m} \partial_{k-m} F. \end{aligned}$$

Since $\binom{n}{m-e_i} + \binom{n}{m} = \binom{n+e_i}{m}$, it turns out that (6.12) also holds for $n + e_i$. ■

Using the Faà di Bruno formula (6.11), we can give in the following lemma a representation of the nonlinear terms of (4.11). Given a modelled distribution $\boldsymbol{u} \in \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ of the form (6.9), set

$$\mathcal{F}(\boldsymbol{u}) := \mathcal{Q}_{<\gamma+\beta_0}\Big(\sum_{\mathfrak{b}\in\mathrm{N}^0} \big(F(\mathfrak{b}^*)\big)^\star(\boldsymbol{u}, D\boldsymbol{u}) \star \mathfrak{b}\Big) = f^\star(\boldsymbol{u}) \star \Xi + g^\star(\boldsymbol{u}, D\boldsymbol{u}).$$

Lemma 35. *Consider a generic tree of the form*

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \bigstar_{j=1}^{b} (\mathcal{I}_{p_j}(\boldsymbol{\sigma}_j))^{\star m_j}$$

such that $|\boldsymbol{\tau}| < \gamma + \beta_0$ and $(p_i, \boldsymbol{\sigma}_i) \neq (p_j, \boldsymbol{\sigma}_j)$ for any $i \neq j$. The $\boldsymbol{\tau}$-component of $\mathcal{F}(\boldsymbol{u}(x))$ is given by

$$\left\{\prod_{j=1}^{b} \frac{u_{\mathcal{I}(\boldsymbol{\sigma}_j)}(x)^{m_j}}{m_j!}\right\}\left\{\frac{D^n}{n!}\prod_{j=1}^{b} \partial_{p_j}^{m_j}\right\} F\big(\mathfrak{b}^*\big)\big((u_k(x))_{|k|_{\mathfrak{s}}\le 1}\big). \tag{6.15}$$

Consequently, if $\boldsymbol{u}$ solves the fixed-point problem (4.11), then for any $\boldsymbol{\tau} \in \mathcal{B}$ with $|\boldsymbol{\tau}| < \gamma - 2$ and any $x \in (0, \frac{t_0}{2}) \times \mathbb{T}^d$, one has

$$u_{\mathcal{I}(\boldsymbol{\tau})}(x) = \frac{1}{S(\boldsymbol{\tau})} F(\boldsymbol{\tau}^*)\big((u_k(x))_{k\in\mathbb{N}^{1+d}}\big), \tag{6.16}$$

where the right-hand side depends only on $u_k(x)$ with $|k|_{\mathfrak{s}} < \gamma$.

Proof. We consider here

$$\boldsymbol{\tau} = \circ^n \star \bigstar_{j=1}^{b} \big(\mathcal{I}(\boldsymbol{\sigma}_j)\big)^{\star m_j};$$

the other cases are proved by similar arguments. The element $\boldsymbol{\tau}$ appears in the term $f^\star(\boldsymbol{u}) \star \Xi$. Inserting the expansion (6.9) into $f^\star(\boldsymbol{u}) \star \Xi$, its $\boldsymbol{\tau}$-component is given by

$$\sum_{\boldsymbol{\tau}_1,\dots,\boldsymbol{\tau}_a;\, \bigstar_{i=1}^{a}\boldsymbol{\tau}_i=\boldsymbol{\tau}} \frac{f^{(a)}(u_0)}{a!} u_{\boldsymbol{\tau}_1}\cdots u_{\boldsymbol{\tau}_a},$$

where $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_a$ are elements of $\mathcal{B}$ of the forms X^k with $k \neq 0$ or $\mathcal{I}(\boldsymbol{\sigma}_j)$, and $u_{X^k} = \frac{u_k}{k!}$ by definition. Rearranging $\boldsymbol{\tau}_1, \dots, \boldsymbol{\tau}_a$ so that duplicate elements are grouped together, the above quantity is reorganized as

$$\sum_{\substack{q_1,\dots,q_r\in\mathbb{N}^{1+d}\setminus\{0\},\, i\neq j\Rightarrow q_i\neq q_j,\\ n_1,\dots,n_r\in\mathbb{N},\, n_1q_1+\cdots+n_rq_r=n}} \times \frac{f^{(n_1+\cdots+n_r+m_1+\cdots+m_b)}(u_0)}{n_1!\cdots n_r! m_1!\cdots m_b!}\left(\frac{u_{q_1}}{q_1!}\right)^{n_1}\cdots\left(\frac{u_{q_r}}{q_r!}\right)^{n_r} u_{\mathcal{I}(\boldsymbol{\sigma}_1)}^{m_1}\cdots u_{\mathcal{I}(\boldsymbol{\sigma}_b)}^{m_b}.$$

Applying the formula (6.11) to the sum over $q_1, \dots, q_r$ and $n_1, \dots, n_r$, the above quantity is equal to

$$\frac{1}{n!\, m_1! \cdots m_b!}\, u^{m_1}_{\mathcal{I}(\boldsymbol{\sigma}_1)} \cdots u^{m_b}_{\mathcal{I}(\boldsymbol{\sigma}_b)} D^n \partial_0^{m_1+\cdots+m_b} f(u_0).$$

This is a particular case of (6.15) with $\mathfrak{b} = \circ$ and $p_j = 0$.

If $\boldsymbol{u}$ solves the fixed-point problem (4.11), then

$$\boldsymbol{u}(x) = \sum_{|k|_{\mathfrak{s}}<\gamma} \frac{u_k(x)}{k!} X^k + \mathcal{Q}_{<\gamma} \mathcal{I}\big(\mathcal{F}(\boldsymbol{u}(x))\big)$$

for any $x \in (0, \frac{t_0}{2}) \times \mathbb{T}^d$. Therefore, the quantity (6.15) should be equal to $u_{\mathcal{I}(\boldsymbol{\tau})}(x)$ for any $\boldsymbol{\tau} \in \mathcal{B}$ with $|\boldsymbol{\tau}| < \gamma - 2$. Assuming

$$u_{\mathcal{I}(\boldsymbol{\sigma}_j)} = F(\boldsymbol{\sigma}_j^*)/S(\boldsymbol{\sigma}_j)$$

inductively, we have

$$u_{\mathcal{I}(\boldsymbol{\tau})} = \frac{1}{S(\boldsymbol{\tau})} F(\boldsymbol{\tau}^*).$$

This concludes the proof. ■

Modelled distributions satisfying identity (6.16) are called "*coherent*" in [17].

Recall $\mathsf{E} \simeq \mathbb{N}^{d+1}$, and define the family of differential operators

$$G \triangleright_{\mathsf{e}} H := G\, \partial_{u_{\mathsf{e}}} H \quad (\mathsf{e} \in \mathsf{E}),$$

acting on smooth functions of (u_0, u_1), with $u_1 = (u_{X_i})_{i=1}^d$. The family $\{\triangleright_{\mathsf{e}}\}_{\mathsf{e}\in\mathsf{E}}$ defines an E-multi-pre-Lie algebra structure.

Proposition 36. *The map F is an E-multi-pre-Lie algebra morphism: for any $\mathsf{e} \in \mathsf{E}$ and any decorated trees $\boldsymbol{\tau}, \boldsymbol{\sigma}$ in $\mathcal{B}$, one has*

$$F\big(\boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}^*\big) = F(\boldsymbol{\tau}^*) \triangleright_{\mathsf{e}} F(\boldsymbol{\sigma}^*). \tag{6.17}$$

Proof. Assume that $\boldsymbol{\sigma}$ is of the form

$$\boldsymbol{\sigma} = \mathfrak{b}^n \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\sigma}_i).$$

Then, by definition,

$$\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma} = \sum_{\ell} \binom{n}{\ell} \mathfrak{b}^{n-\ell} \star \mathcal{I}_{\mathsf{e}-\ell}(\boldsymbol{\tau}) \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\sigma}_i)$$
$$+ \sum_{i=1}^{a} \mathfrak{b}^n \star \mathcal{I}_{p_i}\big(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}_i\big) \star \mathop{\bigstar}_{j;j\neq i} \mathcal{I}_{p_j}(\boldsymbol{\sigma}_j).$$

Hence,

$$F\big(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright} \boldsymbol{\sigma}^*\big) = \sum_{\ell} \binom{n}{\ell} F(\boldsymbol{\tau}^*) \left\{ \prod_{i=1}^{a} F(\boldsymbol{\sigma}_i^*) \right\} D^{n-\ell} \partial_{\mathrm{e}-\ell} \left\{ \prod_{i=1}^{a} \partial_{p_i} \right\} F(\mathfrak{b}^*)$$
$$+ \sum_{i=1}^{a} F(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright} \boldsymbol{\sigma}_i^*) \Big\{ \prod_{j;j\neq i} F(\boldsymbol{\sigma}_i^*) \Big\} D^n \left\{ \prod_{i=1}^{a} \partial_{p_i} \right\} F(\mathfrak{b}^*).$$

On the other hand, by Leibniz rule,

$$F(\boldsymbol{\tau}^*) \rhd_{\mathrm{e}} F(\boldsymbol{\sigma}^*) = F(\boldsymbol{\tau}^*) \left\{ \prod_{i=1}^{a} F(\boldsymbol{\sigma}_i^*) \right\} \partial_{\mathrm{e}} D^n \left\{ \prod_{i=1}^{a} \partial_{p_i} \right\} F(\mathfrak{b}^*)$$
$$+ F(\boldsymbol{\tau}^*) \sum_{i=1}^{a} \partial_{\mathrm{e}} F(\boldsymbol{\sigma}_i^*) \Big\{ \prod_{j;j\neq i} F(\boldsymbol{\sigma}_i^*) \Big\} D^n \left\{ \prod_{i=1}^{a} \partial_{p_i} \right\} F(\mathfrak{b}^*).$$

The first terms in the expansions of $F(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright} \boldsymbol{\sigma}^*)$ and $F(\boldsymbol{\tau}^*) \rhd_{\mathrm{e}} F(\boldsymbol{\sigma}^*)$ coincide because of the identity (6.12). The second terms turn out to coincide if (6.17) holds for $\boldsymbol{\tau}^*$ and $\boldsymbol{\sigma}_i^*$, which leads an induction on the number of edges contained in $\boldsymbol{\sigma}^*$. ■

Assumption D1 is a necessary condition for the basis $\mathcal{B}$. The next assumption means that $\mathcal{B}$ is sufficiently large to describe all terms on the right-hand side of (6.2).

Assumption D2. One has $F(\boldsymbol{\tau}^*) = 0$ for any $\boldsymbol{\tau} \in \mathcal{V} \setminus \mathcal{B}$.

In particular, Assumption D2 holds if $\mathcal{B}$ contains all trees strongly conforming to the rule as in Section 9.1. Indeed, if $\boldsymbol{\tau}$ is not strongly conforming and does not have any node with $\bullet^{\cdot,\alpha}$ decoration, then $\boldsymbol{\tau}$ have an edge $\mathcal{I}_p$ with $|p|_{\mathfrak{s}} \geq 2$ or have a node with at least three leaving edges $\mathcal{I}_p$ with $|p|_{\mathfrak{s}} = 1$. Since $F(\bullet^*)$ is at most quadratic with respect to u_1, we have $F(\boldsymbol{\tau}^*) = 0$. We define

$$\Upsilon := F|_{U^*}.$$

By Assumption D2, we can conclude that Υ is an E-multi-pre-Lie algebra morphism on the E-multi-pre-Lie algebra $(U^*, \{\overset{\mathrm{e}}{\curvearrowright}_{\mathfrak{b}}\}_{\mathrm{e}\in\mathsf{E}})$.

Proposition 37. *Under Assumption D2, the map Υ is an E-multi-pre-Lie algebra morphism: for any $\mathrm{e} \in \mathsf{E}$ and any decorated trees $\boldsymbol{\tau}$, $\boldsymbol{\sigma}$ in $\mathcal{B}$, one has*

$$\Upsilon\big(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright}_{\mathfrak{b}} \boldsymbol{\sigma}^*\big) = \Upsilon(\boldsymbol{\tau}^*) \rhd_{\mathrm{e}} \Upsilon(\boldsymbol{\sigma}^*).$$

Proof. Since $\Upsilon \circ \pi_{U^*} = F \circ \pi_{U^*} = F$,

$$\Upsilon\big(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright}_{\mathfrak{b}} \boldsymbol{\sigma}^*\big) = F\big(\boldsymbol{\tau}^* \overset{\mathrm{e}}{\curvearrowright} \boldsymbol{\sigma}^*\big) = F(\boldsymbol{\tau}^*) \rhd_{\mathrm{e}} F(\boldsymbol{\sigma}^*) = \Upsilon(\boldsymbol{\tau}^*) \rhd_{\mathrm{e}} \Upsilon(\boldsymbol{\sigma}^*). \qquad ■$$

The next proposition is proved by an induction similar to the induction used in the proof of Proposition 36, noting that D_i satisfies Leibniz rule.

Proposition 38. *Under Assumption* D2*, for any* $i \in \{0, 1, \dots, d\}$ *and* $\boldsymbol{\tau} \in \mathcal{V}$*, one has*

$$F(\uparrow_i \boldsymbol{\tau}^*) = D_i F(\boldsymbol{\tau}^*). \tag{6.18}$$

And for $\boldsymbol{\tau} \in \mathcal{B}$*,*

$$\Upsilon(\uparrow_i \boldsymbol{\tau}^*) = D_i \Upsilon(\boldsymbol{\tau}^*). \tag{6.19}$$

Proof. Assume that $\boldsymbol{\tau}$ is of the form

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \mathop{\bigstar}_{k=1}^{a} \mathcal{I}_{p_k}(\boldsymbol{\tau}_k).$$

Then, by definition,

$$\uparrow_i \tau = \mathfrak{b}^{n+e_i} \star \mathop{\bigstar}_{k=1}^{a} \mathcal{I}_{p_k}(\boldsymbol{\tau}_k) + \sum_{k=1}^{a} \mathfrak{b}^n \star \mathcal{I}_{p_k}(\uparrow_i \boldsymbol{\tau}_k) \star \mathop{\bigstar}_{j;j\neq k} \mathcal{I}_{p_j}(\boldsymbol{\tau}_j).$$

Hence,

$$\begin{aligned} F(\uparrow_i \boldsymbol{\tau}^*) = &\left\{\prod_{k=1}^{a} F(\boldsymbol{\tau}_k^*)\right\} D^{n+e_i} \left\{\prod_{k=1}^{a} \partial_{p_k}\right\} F(\mathfrak{b}^*) \\ &+ \sum_{k=1}^{a} F(\uparrow_i \boldsymbol{\tau}_k^*) \Big\{\prod_{j;j\neq k} F(\boldsymbol{\tau}_j^*)\Big\} D^n \left\{\prod_{k=1}^{a} \partial_{p_k}\right\} F(\mathfrak{b}^*). \end{aligned}$$

If (6.18) holds for $\boldsymbol{\tau}_i$, then the above quantity is equal to $D_i F(\boldsymbol{\tau}^*)$ by Leibniz rule. Thus, the proof is reduced to an induction on the number of edges contained in $\boldsymbol{\tau}$. One also obtains (6.19) by Assumption D2. ∎

6.3. Renormalization structure over a multi-pre-Lie algebra

We now come to the main result of [17], giving a dynamical meaning to the renormalization operations on models associated with elements $k \in G_{\text{ad}}^-$ of the renormalization group and more generally to elements $k \in G^-$. We keep working on the example of the generalized KPZ equation.

In Theorem 47 in Section 9, we show that one can choose U stable under all the splitting maps $\upharpoonleft\!\downharpoonright_{\mathsf{e}}$, that is,

$$\upharpoonleft\!\downharpoonright_{\mathsf{e}}(U) \subset U \otimes U$$

for any $\mathsf{e} \in \mathsf{E}$. The restricted map

$$\upharpoonleft\!\downharpoonright_{\mathsf{e}}|_U : U \to U \otimes U$$

is then the dual of the map $\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}$ for any $\mathsf{e} \in \mathsf{E}$. The following assumption is thus to be understood as a constraint on which renormalization schemes δ can be used.

Assumption D3. (a) For any $\mathsf{e} \in \mathsf{E}$, the space U is stable under $\uparrow\!\!\mid\!\!\blacktriangleright_{\mathsf{e}}$, and one has

$$\big(\mathrm{Id} \otimes (\uparrow\!\!\mid\!\!\blacktriangleright_{\mathsf{e}}|_U)\big)\delta = \mathcal{M}^{(13)}\big(\delta \otimes \delta\big) \uparrow\!\!\mid\!\!\blacktriangleright_{\mathsf{e}}|_U \tag{6.20}$$

and

$$\delta \circ \downarrow_i = \big(\mathrm{Id} \otimes \downarrow_i\big)\delta. \tag{6.21}$$

(b) When $\boldsymbol{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}}$ is an element of $\mathcal{B}$ without $\bullet^{\cdot,\alpha}$ decorations, then, for any subforest φ of τ, $\mathfrak{n}_\varphi : N_\varphi \to \mathbb{N}^{d+1}$ with $\mathfrak{n}_\varphi \leq \mathfrak{n}$, and $\mathfrak{e}'_{\partial\varphi} : \partial\varphi \to \mathbb{N}^{d+1}$, the tree

$$(\tau/^{\mathrm{red}}\varphi)_{\mathfrak{e}+\mathfrak{e}'_{\partial\varphi}}^{[\mathfrak{n}-\mathfrak{n}_\varphi]_\varphi,\mathfrak{o}(\mathfrak{n}_\varphi+\pi\mathfrak{e}'_{\partial\varphi},\mathfrak{e})} \tag{6.22}$$

is also contained in $\mathcal{B}$ – see Section 9.2 for the notation.

Assumptions D1, D2, and D3 are jointly called Assumption D. Identity (6.20) is the E-multi-pre-Lie version of the compatibility condition (5.6) between the splitting map Δ of a regularity structure and a renormalization splitting δ. Recall that any character k of U^- defines a linear map $\tilde{k} = (k \otimes \mathrm{Id})\delta : U \to U$. Denote by $\tilde{k}^* : U^* \to U^*$ the dual map of $\tilde{k}$ under the pairing (6.3). Anticipating over Section 9, say here that $\bullet^{0,\alpha}$ is used to denote the result of extracting from a decorated tree $\boldsymbol{\tau}$ the entire tree but keeping track of the homogeneity $\alpha = |\boldsymbol{\tau}|$ of the tree that was removed. Using the duality relation defining $\tilde{k}^*$ and the definition of $\tilde{k}$, we see that

$$\tilde{k}^*(\circ^*) = \circ^*, \quad \tilde{k}^*(\bullet^*) = \bullet^*,$$

and

$$\tilde{k}^*\big((\bullet^{0,\alpha})^*\big) = 0$$

for $\alpha > 0$, and

$$\tilde{k}^*\big((\bullet^{0,\alpha})^*\big) = \sum_{\boldsymbol{\tau}\in\mathcal{B},|\boldsymbol{\tau}|=\alpha} \frac{k(\boldsymbol{\tau})}{S(\boldsymbol{\tau})}\,\boldsymbol{\tau}^*$$

for $\alpha < 0$. The following result is part of Proposition 4.18 in Bruned, Chandra, Chevyrev, and Hairer's work [17]. It is the reason why we insisted on making a difference between U and U^* to emphasize the dual action of $\tilde{k}$.

Proposition 39. *Under the compatibility Assumption D3, given any character k on U^-, the map $\tilde{k}^*$ is an E-multi-pre-Lie morphism: for any edge-type $\mathsf{e} \in \mathsf{E}$, and any $\boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathcal{B}$, one has*

$$\tilde{k}^*(\boldsymbol{\tau}^*) \overset{\mathsf{e}}{\curvearrowright}_{\flat} \tilde{k}^*(\boldsymbol{\sigma}^*) = \tilde{k}^*(\boldsymbol{\tau}^* \overset{\mathsf{e}}{\curvearrowright}_{\flat} \boldsymbol{\sigma}^*)$$

and

$$\tilde{k}^* \circ \uparrow_i = \uparrow_i \circ \tilde{k}^* \quad (1 \leq i \leq d). \tag{6.23}$$

Proof. We prove the dual identities writing

$$\Uparrow_{\mathfrak{e}} \circ \tilde{k} = (k \otimes \Uparrow_{\mathfrak{e}})\delta \overset{(6.20)}{=} \big(k \otimes \mathrm{Id} \otimes \mathrm{Id}\big)\mathcal{M}^{(13)}(\delta \otimes \delta) \Uparrow_{\mathfrak{e}}$$
$$= \big((k \otimes \mathrm{Id})\delta \otimes (k \otimes \mathrm{Id})\delta\big) \Uparrow_{\mathfrak{e}} = \big(\tilde{k} \otimes \tilde{k}\big) \Uparrow_{\mathfrak{e}}$$

and

$$\downarrow_i \circ \tilde{k} = (k \otimes \downarrow_i)\delta \overset{(6.21)}{=} (k \otimes \mathrm{Id})\delta \circ \downarrow_i = \tilde{k} \circ \downarrow_i .$$ ■

Pick a character k on U^-. For primitive trees $\mathfrak{b} \in \mathsf{N}^0$, define

$$F^{(k)}(\mathfrak{b}^*) := F\big(\tilde{k}^*(\mathfrak{b}^*)\big), \tag{6.24}$$

so we have

$$F^{(k)}(\circ^*) = f(\mathsf{u}_0), \quad F^{(k)}(\bullet^*) = g(\mathsf{u}_0, \mathsf{u}_1), \quad F^{(k)}\big((\bullet^{0,\alpha})^*\big)$$
$$= \mathbf{1}_{\alpha<0} \sum_{\boldsymbol{\tau} \in \mathcal{B} \cap U_\alpha} \frac{k(\boldsymbol{\tau})}{S(\boldsymbol{\tau})} F(\boldsymbol{\tau}^*).$$

By Lemma 35, $F^{(k)}((\bullet^{0,\alpha})^*)$ is also a function of only u_0 and u_1. For a tree

$$\boldsymbol{\tau} = \mathfrak{b}^n \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i),$$

define inductively the functions of (u_0, u_1)

$$F^{(k)}(\boldsymbol{\tau}^*)(\mathsf{u}) := \left\{\prod_{i=1}^{a} F^{(k)}(\boldsymbol{\tau}_i^*)(u_0, u_1)\right\}\left\{D^n \prod_{i=1}^{a} \partial_{p_i}\right\} F^{(k)}(\mathfrak{b}^*)(\mathsf{u}), \tag{6.25}$$

similarly to (6.10). Note that $F^{(k)}$ can be defined for all elements $\boldsymbol{\tau}$ in $\mathcal{V}$. The map $F^{(k)}$ is an E-multi-pre-Lie morphism with respect to $\{\overset{\mathfrak{e}}{\curvearrowright}\}_{\mathfrak{e}\in\mathsf{E}}$ by the same proof as that of Proposition 36. We prove the following proposition to ensure that $F^{(k)}(\boldsymbol{\tau}^*) = 0$ for any $\boldsymbol{\tau} \in \mathcal{V} \setminus \mathcal{B}$. The notations are all defined in Section 9.2; the reader can skip it now and come back to it later.

Proposition 40. *Let k be a character of U^-. For any $\boldsymbol{\tau} \in \mathcal{V}$, the function $F^{(k)}(\boldsymbol{\tau}^*)$ is represented as a linear combination of the functions $F(\boldsymbol{\sigma}^*)$, where $\boldsymbol{\sigma} = \sigma_{\mathfrak{e}}^{\mathfrak{n}}$ runs over all elements of $\mathcal{V}$ without $\bullet^{\cdot,\alpha}$ decorations and such that $\boldsymbol{\tau}$ is obtained by contracting $\boldsymbol{\sigma}$ by its subforest φ with maps $\mathfrak{n}_\varphi$ and $\mathfrak{e}'_{\partial\varphi}$ as the form (6.22).*

Proof. The proof is an induction on the number of edges contained in $\boldsymbol{\tau}$. It is sufficient to consider $\boldsymbol{\tau}$ of the form

$$\boldsymbol{\tau} = \bullet^{n,\alpha} \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i).$$

Assume that the result holds for each $\boldsymbol{\tau}_i$. Then, by definition, $F^{(k)}(\boldsymbol{\tau}^*)$ is a linear combination of the functions of the form

$$\left\{\prod_{i=1}^{a} F(\boldsymbol{\sigma}_i^*)\right\}\left\{D^n \prod_{i=1}^{a} \partial_{p_i}\right\} F(\boldsymbol{\eta}^*),$$

where $\boldsymbol{\sigma}_i$ is an element of $\mathcal{V}$ without ${\color{red}\bullet}^{\cdot,\alpha}$ decorations such that $\boldsymbol{\tau}_i$ is obtained as a contraction of $\boldsymbol{\sigma}_i$, and $\boldsymbol{\eta} \in \mathcal{B} \cap U_\alpha$. By the formula $D^n \partial_k F = \sum_{\ell \in \mathbb{N}^{1+d}} (-1)^\ell \binom{n}{\ell} \partial_{k-\ell} D^{n-\ell} F$ obtained similarly to (6.12) and by Proposition 38, the above function is a linear combination of the functions

$$\left\{\prod_{i=1}^{a} F(\boldsymbol{\sigma}_i^*)\right\}\left\{\prod_{i=1}^{a} \partial_{q_i}\right\} F(\boldsymbol{\mu}^*),$$

where $q_i \le p_i$ and $\boldsymbol{\mu} \in \mathcal{B}$ is a tree appearing in the expansion of $\uparrow^m \boldsymbol{\eta} := (\prod_{i=0}^d \uparrow_i^{m_i})\boldsymbol{\eta}$ for some $m \in \mathbb{N}^{1+d}$. By an argument similar to the proof of Proposition 36, we can show that the above function is equal to

$$F\Big(\sum_{v_1,\dots,v_a \in N_\mu} \big(\boldsymbol{\sigma}_1 \overset{q_1}{\curvearrowright}_{(v_1)} (\boldsymbol{\sigma}_2 \overset{q_2}{\curvearrowright}_{(v_2)} \cdots (\boldsymbol{\sigma}_a \overset{q_a}{\curvearrowright}_{(v_a)} \boldsymbol{\mu}) \cdots)\big)^*\Big).$$

The trees inside F produce $\boldsymbol{\tau}$ when we contract $\boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_a, \boldsymbol{\mu}$ as in the explicit formula of D^- in Section 9.2. ■

As a result of Proposition 40, under Assumptions D2 and D3, we have that $F^{(k)}(\boldsymbol{\tau}^*) = 0$ for any $\boldsymbol{\tau} \in \mathcal{V} \setminus \mathcal{B}$, since $\boldsymbol{\tau}$ is an element of $\mathcal{B}$ if and only if $\boldsymbol{\tau}$ is a contraction of a tree in $\mathcal{B}$ without ${\color{red}\bullet}^{\cdot,\alpha}$ decoration. This ensures that the map

$$\Upsilon^{(k)} := F^{(k)}|_{U^*}$$

is an E-multi-pre-Lie morphism on U^* with respect to $\{\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}\}_{\mathsf{e}\in E}$. Moreover, denoting by S the subspace of T spanned by $\{\bullet^n\}_{n\in\mathbb{N}^{d+1}} \cup \mathcal{I}(\mathcal{B})$, we have that, for any $\mathfrak{b} \in \mathsf{N}^0$ and any function $\boldsymbol{u} : \mathbb{R}^{d+1} \to S$, the function

$$\big(F^{(k)}(\mathfrak{b}^*)\big)^\star(\boldsymbol{u}, D\boldsymbol{u}) \star \mathfrak{b} : \mathbb{R}^{d+1} \to V$$

is actually T-valued.

Corollary 41. *Under Assumption* D, *one has* $\Upsilon \circ \tilde{k}^* = \Upsilon^{(k)}$ *for all* $k \in G^-$.

Proof. It follows from Propositions 36 and 39 that the map $\Upsilon \circ \tilde{k}^*$ is an ($\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}$ vs $\triangleright_{\mathsf{e}}$) E-multi-pre-Lie morphism. Because of Proposition 40, the map $\Upsilon^{(k)}$ is also an ($\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}$ vs $\triangleright_{\mathsf{e}}$) E-multi-pre-Lie morphism. Hence, it is sufficient to show that they are equal on the generators of U^*, that is,

$$F\big(\tilde{k}^*(\mathfrak{b}^n)^*\big) = F^{(k)}\big((\mathfrak{b}^n)^*\big)$$

for any $(\mathfrak{b}, n) \in \mathfrak{T}_n \times \mathbb{N}^{d+1}$. The case $n = 0$ is given by definition (6.24). For $n = (n_i)_{i=0}^d \in \mathbb{N}^{d+1}$, by writing $\uparrow^n := \prod_{i=0}^d \uparrow_i^{n_i}$, we have

$$
\begin{aligned}
F\big(\tilde{k}^*(\mathfrak{b}^n)^*\big) &= F\big(\tilde{k}^*(\uparrow^n \mathfrak{b}^*)\big) \overset{(6.23)}{=} F\big(\uparrow^n (\tilde{k}^*\mathfrak{b}^*)\big) \overset{(6.19)}{=} D^n F\big(\tilde{k}^*\mathfrak{b}^*\big) \overset{(6.24)}{=} D^n F^{(k)}(\mathfrak{b}^*) \\
&= F^{(k)}\big((\mathfrak{b}^n)^*\big),
\end{aligned}
$$

using (6.19) and (6.23) in the last equality. ■

Similarly to the definition of $\mathcal{F}(\boldsymbol{u})$, given a modelled distribution $\boldsymbol{u} \in \mathcal{D}^{\gamma,\eta}(T, \mathrm{g})$ with $\gamma \in (-\beta_0, 2)$ such that $(0, \gamma + \beta_0) \cap A = \emptyset$, set

$$
\mathcal{F}^{(k)}(\boldsymbol{u}) := \mathcal{Q}_{\le 0}\Big(\sum_{\mathfrak{b}\in\mathrm{N}^0} \big(F^{(k)}(\mathfrak{b}^*)\big)^{\star}(\boldsymbol{u}, D\boldsymbol{u}) \star \mathfrak{b}\Big). \tag{6.26}
$$

Note the appearance in (6.26) of a number of symbols $\color{red}{\bullet}^{0,\alpha}$, with $\alpha < 0$, that have no counterpart in $\mathcal{F}(\boldsymbol{u})$.

Lemma 42. *If $\boldsymbol{u}$ is a solution of equation (4.11), then*

$$
\tilde{k}\big(\mathcal{F}(\boldsymbol{u})\big) = \mathcal{F}^{(k)}\big(\tilde{k}(\boldsymbol{u})\big) \tag{6.27}
$$

on the domain $(0, \frac{t_0}{2}) \times \mathbb{T}^d$.

Proof. Let $\boldsymbol{u}$ be evaluated at the fixed x in the domain $(0, \frac{t_0}{2}) \times \mathbb{T}^d$. Lemma 35 implies that

$$
\mathcal{F}(\boldsymbol{u}) = \sum_{|\boldsymbol{\tau}|<\gamma+\beta_0} \frac{\Upsilon(\boldsymbol{\tau}^*)\big((u_k)_{k\in\mathbb{N}^{1+d}}\big)}{S(\boldsymbol{\tau})}\boldsymbol{\tau},
$$

or, equivalently, $\langle\!\langle \mathcal{F}(\boldsymbol{u}), \boldsymbol{\tau}^* \rangle\!\rangle = \Upsilon(\boldsymbol{\tau}^*)$ for any $\boldsymbol{\tau} \in \mathcal{B}$ with $|\boldsymbol{\tau}| < \gamma + \beta_0$. Noting that $\tilde{k}$ and $\tilde{k}^*$ preserve the grading of T, as a consequence of the compatibility condition (5.3), we have

$$
\langle\!\langle \tilde{k}(\mathcal{F}(\boldsymbol{u})), \boldsymbol{\tau}^* \rangle\!\rangle = \langle\!\langle \mathcal{F}(\boldsymbol{u}), \tilde{k}^*(\boldsymbol{\tau}^*) \rangle\!\rangle = \Upsilon\big(\tilde{k}^*(\boldsymbol{\tau}^*)\big) = \Upsilon^{(k)}(\boldsymbol{\tau}^*)
$$

for any $\boldsymbol{\tau} \in \mathcal{B}$ with $|\boldsymbol{\tau}| < \gamma + \beta_0$, or, equivalently,

$$
\tilde{k}\big(\mathcal{F}(\boldsymbol{u})\big) = \sum_{|\boldsymbol{\tau}|<\gamma+\beta_0} \frac{\Upsilon^{(k)}(\boldsymbol{\tau}^*)\big((u_k)_{k\in\mathbb{N}^{1+d}}\big)}{S(\boldsymbol{\tau})}\boldsymbol{\tau}.
$$

Next, we consider the $\boldsymbol{\tau}$-component of $\mathcal{F}^{(k)}(\tilde{k}(\boldsymbol{u}))$ by an argument similar to Lemma 35. Note that Assumption C yields $\tilde{k}(X^n) = X^n$, and $\tilde{k}(\mathcal{I}(\boldsymbol{\tau})) = \mathcal{I}(\tilde{k}(\boldsymbol{\tau}))$. Hence,

$$
\tilde{k}(\boldsymbol{u}) = \tilde{k}\Big\{\sum_{|k|_{\mathfrak{s}}<\gamma} \frac{u_k}{k!}X^k + \mathcal{Q}_{<\gamma}\mathcal{I}\big(\mathcal{F}(\boldsymbol{u})\big)\Big\} = \sum_{|k|_{\mathfrak{s}}<\gamma} \frac{u_k}{k!}X^k + \sum_{\boldsymbol{\tau}} \frac{\Upsilon^{(k)}(\boldsymbol{\tau}^*)}{S(\boldsymbol{\tau})}\mathcal{I}(\boldsymbol{\tau}).
$$

Thus, by an argument similar to the former half part of Lemma 35, we see that the $\boldsymbol{\tau}$-component of $\mathcal{F}^{(k)}(\tilde{k}(\boldsymbol{u}))$ is equal to (6.15), where $u_{\mathcal{I}(\boldsymbol{\sigma}_j)}$ is replaced by $\frac{\Upsilon^{(k)}(\boldsymbol{\sigma}_j^*)}{S(\boldsymbol{\sigma}_j)}$ and $F(\mathfrak{b}^*)$ is replaced by $\Upsilon^{(k)}(\mathfrak{b}^*)$. By definition of $\Upsilon^{(k)}(\boldsymbol{\tau}^*)$, we see that the $\boldsymbol{\tau}$-component of $\mathcal{F}^{(k)}(\tilde{k}(\boldsymbol{u}))$ is equal to $\frac{\Upsilon^{(k)}(\boldsymbol{\tau}^*)}{S(\boldsymbol{\tau})}$. Therefore,

$$\langle\!\langle \tilde{k}(\mathcal{F}(\boldsymbol{u})), \boldsymbol{\tau}^* \rangle\!\rangle = \langle\!\langle \mathcal{F}^{(k)}(\tilde{k}(\boldsymbol{u})), \boldsymbol{\tau}^* \rangle\!\rangle$$

for any $\boldsymbol{\tau} \in \mathcal{B}$ with $|\boldsymbol{\tau}| < \gamma + \beta_0$. ■

The next statement provides a dynamical picture of the renormalization operation on models. As its proof will make it clear, it is a consequence of identity (6.27) and Theorem 28, giving in particular the reconstruction operator of a renormalized smooth model in terms of the unrenormalized smooth model, together with the multiplicativity property of the canonical model associated with a smooth noise.

Theorem 43. *Assume Assumption* D. *Let ζ be a smooth noise with canonical model $\mathsf{M}^\zeta = (\Pi^\zeta, g^\zeta)$. Given a character $k \in G_{ad}^-$, denote by ${}^k\mathsf{M}^\zeta = (\mathsf{g}^\zeta \circ \tilde{k}^+, \Pi^\zeta \circ \tilde{k})$ its associated renormalized* **K***-admissible model – see Theorem 28. Pick $\eta \in (0, \beta_0 + 2]$ and $\gamma \in (-\beta_0, 2)$. Given an initial condition $v \in \mathcal{C}^\eta(\mathbb{T}^d)$, let $\boldsymbol{u}^{(k)} \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathsf{g}^\zeta \circ \tilde{k}^+)$ stand for the solution on $(0, t_0)$ to the equation*

$$\boldsymbol{u}^{(k)} = P_\gamma v + \mathscr{P}_{t_0}^{{}^k\mathsf{M}^\zeta}\big(f^\star\big(\boldsymbol{u}^{(k)}\big)\Xi + g^\star\big(\boldsymbol{u}^{(k)}, D\boldsymbol{u}^{(k)}\big)\big).$$

Then,

$$u^{(k)} := \mathbf{R}^{{}^k\mathsf{M}^\zeta}(C_{t_0}\boldsymbol{u}^{(k)})$$

is the solution on $(0, \frac{t_0}{2})$ to the well-posed equation

$$\begin{aligned}\big(\partial_{x_0} - \Delta_{x'} + 1\big)u^{(k)} = f\big(u^{(k)}\big)\zeta + g\big(u^{(k)}, \partial_{x'}u^{(k)}\big) \\ + \sum_{\tau \in \mathcal{B}, |\tau| < 0} \frac{k(\boldsymbol{\tau})}{S(\boldsymbol{\tau})} F(\boldsymbol{\tau}^*)\big(u^{(k)}, \partial_{x'}u^{(k)}\big)\end{aligned}$$

started from v.

Proof. The proof is similar to the proof of Proposition 25. The function $u^{(k)}$ satisfies the equation

$$u^{(k)}(x) = Pv(x) + \int_{(0,x_0)\times\mathbb{T}^d} P(x,y)\mathbf{R}^{{}^k\mathsf{M}^\zeta} C_{t_0}\big(f^\star\big(\boldsymbol{u}^{(k)}\big)\Xi + g^\star\big(\boldsymbol{u}^{(k)}, D\boldsymbol{u}^{(k)}\big)\big)(y)dy.$$

Since ${}^k\mathsf{M}^\zeta$ is a smooth model, one has

$$\mathbf{R}^{{}^k\mathsf{M}^\zeta}(\boldsymbol{w})(x) = \Pi_x^\zeta\big(\tilde{k}\big(\boldsymbol{w}(x)\big)\big)(x)$$

for any modelled distribution $\boldsymbol{w} \in \mathscr{D}^\alpha(T, \mathrm{g}^\zeta \circ \widetilde{k}^+)$ with $\alpha > 0$. Applying Lemma 42 to $\boldsymbol{u}^{(k)}$, one has

$$\begin{aligned}\mathbf{R}^{k\mathsf{M}^\zeta} C_{t_0}\big[f^\star(\boldsymbol{u}^{(k)})\Xi + g^\star\big(\boldsymbol{u}^{(k)}, D\boldsymbol{u}^{(k)}\big)\big](x) \\ = \Pi_x^\zeta\big[\widetilde{k}\big(\mathscr{F}\big(\boldsymbol{u}^{(k)}(x)\big)\big)\big](x) = \Pi_x^\zeta\big[\mathscr{F}^{(k)}\big(\widetilde{k}\big(\boldsymbol{u}^{(k)}(x)\big)\big)\big](x)\end{aligned}$$

for any $x \in (0, \frac{t_0}{2}) \times \mathbb{T}^d$. We see from the definition of the $F^{(k)}$ that the term $\mathscr{F}^{(k)}(\boldsymbol{w})$ is a sum of functions of the form

$$H^\star(\boldsymbol{w})R^\star(D\boldsymbol{w})$$

for smooth functions $H : \mathbb{R} \to \mathbb{R}$ and polynomials R that are at most quadratic. We now use the fact that since the map Π^ζ is *multiplicative*, so is its associated reconstruction operator. The latter has value $\Pi_x^\zeta(\cdot)(x)$ at point x, so we have

$$\Pi_x^\zeta\big[H^\star\big(\boldsymbol{w}(x)\big)R^\star\big(D\boldsymbol{w}(x)\big)\big](x) = H\big[\big(\Pi_x^\zeta \boldsymbol{w}(x)\big)(x)\big]R\big[\big(\Pi_x^\zeta D\boldsymbol{w}(x)\big)(x)\big],$$

with $\boldsymbol{w}(x) = \widetilde{k}(\boldsymbol{u}^{(k)}(x))$. Since

$$\big(\Pi_x^\zeta \boldsymbol{w}(x)\big)(x) = \mathbf{R}^{k\mathsf{M}^\zeta}\big(C_{t_0}\boldsymbol{u}^{(k)}\big)(x) = u^{(k)}(x)$$

and

$$\big(\Pi_x^\zeta D\boldsymbol{w}(x)\big)(x) = \mathbf{R}^{k\mathsf{M}^\zeta}\big(C_{t_0}D\boldsymbol{u}^{(k)}\big)(x) = \partial_{x'}u^{(k)}(x)$$

on the domain $(0, \frac{t_0}{2}) \times \mathbb{T}^d$, we have in the end

$$\Pi_x^\zeta\big[\mathscr{F}^{(k)}\big(\widetilde{k}\big(\boldsymbol{u}^{(k)}(x)\big)\big)\big](x) = \mathscr{F}^{(k)}\big(u^{(k)}, \partial_{x'}u^{(k)}\big)(x). \qquad \blacksquare$$

Remark. The preceding proof underlines the fundamental role played by the multiplicative property of the centered naive interpretation operators Π_x^ζ. The canonical smooth model Π^ζ is not the only multiplicative model that one can associate with a smooth noise ζ. The class of models associated with "preparation maps" introduced by Bruned in [16] provides a general setting where to obtain the renormalized equation for a class of renormalization procedures, including the procedure implemented here [6].

7. The BHZ character

Among all the characters k on U^- that can be used to build a renormalization map $\widetilde{k}$, Bruned, Hairer, and Zambotti proved in [20] that there is a unique character k whose associated random model is centered and translation invariant, in a probabilistic sense, when the smooth noise ζ in the preceding section is random, centered and translation invariant. We describe it in this section and name it "BHZ character", after the initials of Bruned, Hairer, and Zambotti. We also call the associated renormalized model the BHZ model.

Arrived at that stage, the only pieces of the story that will be missing to have a complete proof of the meta-theorems from Section 1 will be a proof of the fact that one can indeed construct some regularity structures satisfying the different assumptions that we put forward in the course of obtaining the above results and a proof of convergence of the BHZ smooth renormalized models. We will tackle the first point in Section 9. The second point is the object of Chandra and Hairer's work [26]; we do not treat it here. We refer the reader to the work [11, 61, 70] for some alternative proofs of the convergence of BHZ renormalized models in situations where the law of the noise satisfies a spectral gap inequality.

We assume throughout this section that we work with regularity and renormalization structures satisfying Assumptions A–C. To have a picture in mind, think of the structures associated with the generalized KPZ equation (6.1). Elements of $T = U$ are thus given by node and edge decorated trees. Denote by $\mathbb{R}[U]$ the commutative algebra generated by U. Recall from Assumption C1 that U^- is an algebra generated by $\mathcal{U}_{<0}$ and a unit $\mathbf{1}_-$, and if one extends first the splitting map $\delta : U \to U^- \otimes U$ into an algebra morphism $\widehat{\delta} : \mathbb{R}[U] \to U^- \otimes \mathbb{R}[U]$, then the splitting map δ^- satisfies

$$\delta^- = (\mathrm{Id} \otimes P_-)\widehat{\delta}|_{U^-}$$

for an algebra morphism projection map $P_- : \mathbb{R}[U] \to U^-$. (See Section 9.3 for the details; this projection map sends in particular $\bullet$ and all the ${\color{red}\bullet}^{0,\alpha}$ to $\mathbf{1}_-$.) Denote by $\hat{\mathbf{1}}_-$ the unit of $\mathbb{R}[U]$, seen as the empty graph. (We use a distinct notation for $\hat{\mathbf{1}}_-$ and $\mathbf{1}_-$ to emphasize that they do not live in the same space.) For basis elements $\boldsymbol{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}}$ and $\boldsymbol{\sigma} = \sigma_{\mathfrak{f}}^{\mathfrak{m}}$ of $\mathbb{R}[U]$, we write

$$\boldsymbol{\sigma} \prec \boldsymbol{\tau}$$

if σ is a strict *subgraph* of τ, or $\sigma = \tau$ and $\mathfrak{m} \leq \mathfrak{n}$ with $\mathfrak{m} \neq \mathfrak{n}$. (Do not get misled by the notation; σ may be a product of disjoint subtrees of τ.) Recall from Section 5.2 the notation $\mathscr{F} = \{\mathcal{I}_p\}_{|p|_{\mathfrak{s}} \leq 1} \cup \{X^n \star\}_{n \in \mathbb{N}^{1+d} \setminus \{0\}}$ for the family of integral operators and multiplication by a non-null monomial. The following assumption describes the properties from the splitting map δ that are relevant here. The renormalization structure built in Section 9 for the generalized KPZ equation satisfies it.

Assumption E. (a) For any $\boldsymbol{\tau} \in \mathcal{U}_{<0}$, one has the splitting formula

$$\delta\boldsymbol{\tau} \in P_-(\boldsymbol{\tau}) \otimes {\color{red}\bullet}^{0,|\tau|} + U_{\boldsymbol{\tau}}^- \otimes U,$$

where

$$U_{\boldsymbol{\tau}}^- := \mathrm{span}\big\{P_-(\boldsymbol{\sigma}) \in U^-;\, \boldsymbol{\sigma} \prec \boldsymbol{\tau}\big\}.$$

(b) For any $\boldsymbol{\tau} \in \mathcal{U}_{<0}$ and $F \in \mathscr{F}$ such that $F(\boldsymbol{\tau}) \in \mathcal{U}_{<0}$, one has

$$\delta\big(F(\boldsymbol{\tau})\big) \in (\mathrm{Id} \otimes F)\delta\boldsymbol{\tau} + P_-\big(F(\boldsymbol{\tau})\big) \otimes {\color{red}\bullet}^{0,|F(\boldsymbol{\tau})|} + \mathfrak{J}_{\boldsymbol{\tau}}^- \otimes U,$$

where $\mathfrak{J}_{\boldsymbol{\tau}}^-$ is the ideal of U^- generated by $\{P_-(F(\boldsymbol{\sigma}));\ F \in \mathscr{F},\ \boldsymbol{\sigma} \in \mathcal{U}_{<0},\ \boldsymbol{\sigma} \prec \boldsymbol{\tau}\}$.

Property (b) is a refinement of the property (5.11) in Assumption C1. Note that, under the definitions of gradings of T and U in Section 9 later, $\bullet^{0,\beta} \in \mathcal{B}_\beta$ but $\bullet^{0,\beta} \in \mathcal{U}_0$ for any $\beta \in \mathbb{R}$. (Basis elements of U with 0-homogeneity are not necessarily unique, unlike in Assumption A1 on concrete regularity structures.) Hence, the above properties are consistent with the definitions of compatible renormalization and regularity structures. Property (b) also ensures the existence of the following map, defined by induction on the order relation $\prec$. Each element $\boldsymbol{\tau}$ of U^- has by definition a unique representative $\boldsymbol{\tau}_U$ in $\mathbb{R}[U]$. Denote by $\mathcal{M}$ the multiplication operator on $\mathbb{R}[U]$ and extend it naturally on $U^- \otimes \mathbb{R}[U]$ setting $\mathcal{M}(\boldsymbol{\tau} \otimes \boldsymbol{\sigma}) = \mathcal{M}(\boldsymbol{\tau}_U \otimes \boldsymbol{\sigma})$; it takes values in $\mathbb{R}[U]$.

Definition. Under Assumption E, the *negative twisted antipode* is an algebra morphism

$$S'_- : U^- \to \mathbb{R}[U]$$

given recursively by $S'_- \mathbf{1}_- = \hat{\mathbf{1}}_-$ and, for every basis element $\boldsymbol{\tau} \in \mathcal{U}_{<0}$ by

$$S'_-\big(P_-(\boldsymbol{\tau})\big) = -\mathcal{M}_-\big(S'_- \otimes \mathrm{Id}\big)\big(\delta\boldsymbol{\tau} - P_-(\boldsymbol{\tau}) \otimes \bullet^{0,|\boldsymbol{\tau}|}\big). \tag{7.1}$$

For the $P_-(\boldsymbol{\tau})$ generating U^- as an algebra for $\boldsymbol{\tau}$ ranging in $\mathcal{U}_{<0}$, identity (7.1) characterizes indeed uniquely an algebra morphism. The intuitive meaning of this recursive definition should be clear. One extracts from $\boldsymbol{\tau}$ all possible subdiverging quantities φ_1 while extracting from φ_1 all its subdiverging quantities, and so on. This formula is close to the Dyson–Salam renormalization formula for the antipode in Hopf algebras [39]; like the latter, it can be rewritten as a sum over forests of diverging sub-forests, as in Zimmermann forest formula. This will not be useful here, and the only thing that matters here is property (7.1). The forest representation is however useful for the analysis of the convergence of renormalized models [26].

Do not be misled by the name of S'_-: this is *not* the antipode of a Hopf algebra structure. Bruned, Hairer, and Zambotti named it so because its defining relation (7.1) looks like the defining relation (B.1) for the antipode in a Hopf algebra.

Recall from Section 6.1 the definition of the naive interpretation operator Π^ζ corresponding to a smooth noise ζ in $\mathbb{R} \times \mathbb{R}^d$. We consider a random smooth noise ζ, invariant by translation and centered. Define the character h^ζ on $\mathbb{R}[U]$ by setting $h^\zeta(\hat{\mathbf{1}}_-) := 1$ and

$$h^\zeta(\boldsymbol{\tau}) := \mathbb{E}\big[\Pi^\zeta \boldsymbol{\tau}\big](0) \tag{7.2}$$

for $\boldsymbol{\tau} \in \mathcal{U}$, and define a character on U^- setting

$$k^\zeta := h^\zeta \circ S'_-.$$

(Keep in mind that S'_- gives back elements of $\mathbb{R}[U]$ and that h^ζ is multiplicative, so k^ζ is multiplicative on U^-.) The associated "*BPHZ* renormalized" interpretation operator ${}^{k^\zeta}\Pi^\zeta$ is defined on U by

$${}^{k^\zeta}\Pi^\zeta \boldsymbol{\tau} = \big(k^\zeta \otimes \Pi^\zeta\big)\delta\boldsymbol{\tau} = \big((h^\zeta \circ S'_-) \otimes \Pi^\zeta\big)\delta\boldsymbol{\tau}.$$

The acronym BPHZ stands for Bogoliubov, Parasiuk, Hepp, and Zimmermann, who made deep contributions to the renormalization problem in quantum field theory. We call *BHZ character*, after Bruned, Hairer, and Zambotti, the character k^ζ on U^-. The reason for introducing the negative twisted antipode operator lies entirely in the following simple computations used in the proof of the next statement claiming that the BPHZ renormalization associated with the BHZ character recenters probabilistically the Π map at all points in spacetime. Its proof is taken from the proof of Theorem 6.17 in Bruned, Hairer, and Zambotti's work [20] on the algebraic renormalization of regularity structures.

Theorem 44. *Let ζ stand for a smooth noise that is centered and translation invariant in law such that $\partial^n\zeta(0)$ has finite moments of any order for any $n \in \mathbb{N}^{d+1}$. We work with compatible regularity and renormalization structures under Assumptions* A–C *and Assumption* E. *The character k^ζ belongs to G^-_{ad}, and one has*

$$\mathbb{E}\big[\big({}^{k^\zeta}\Pi^\zeta\boldsymbol{\tau}\big)(x)\big] = 0 \tag{7.3}$$

for any $\boldsymbol{\tau} \in \mathcal{U}_{<0}$ and $x \in \mathbb{R}\times\mathbb{R}^d$.

Proof. First, we show that $k^\zeta \in G^-_{\mathrm{ad}}$. Let $F \in \mathscr{F}$ and $\boldsymbol{\tau} \in \mathcal{U}_{<0}$ be such that $F\boldsymbol{\tau} \in \mathcal{U}_{<0}$. If $F = X^n\star$, then

$$h^\zeta(F\boldsymbol{\tau}) = \mathbb{E}[\Pi^\zeta X^n \star \boldsymbol{\tau}](0) = x^n\mathbb{E}[\Pi^\zeta\boldsymbol{\tau}(x)]|_{x=0} = 0.$$

Next, consider $F = \mathcal{I}_p$. Recall from Section 3.1 that we defined the operator $\mathbf{K}$ so that

$$\int_{\mathbb{R}\times\mathbb{R}^d} y^p K(x,y)\,dy = 0$$

for all $p \in \mathbb{N}\times\mathbb{N}^d$ such that $|p|_{\mathfrak{s}} < N$ for fixed $N \in \mathbb{N}$. Now, pick $N = 2$. Note that

$$\mathbb{E}[\Pi^\zeta\boldsymbol{\tau}](x) = \mathbb{E}[\Pi^\zeta\tau_e^{\mathfrak{n}}](x)$$

is a polynomial of x with degree $\sum_{v\in N_\tau}|\mathfrak{n}(v)|_{\mathfrak{s}}$, since $K(x,y)$ depends on $x-y$ only. We use it here to have

$$h^\zeta\big(\mathcal{I}_p\boldsymbol{\tau}\big) = \int_{\mathbb{R}\times\mathbb{R}^d}\partial_x^p K(x,y)\mathbb{E}\big[\Pi^\zeta\boldsymbol{\tau}(y)\big] = 0,$$

since $\sum_{v\in N_\tau}|\mathfrak{n}(v)|_{\mathfrak{s}} \le 1$ (otherwise $|\boldsymbol{\tau}| \ge \beta_0 + 2 > 0$). Hence, $h^\zeta(F\boldsymbol{\tau}) = 0$ for all $F \in \mathscr{F}$. Since Assumption E guarantees that we have

$$S'_-(F\boldsymbol{\tau}) \in -\mathcal{M}_-\big(S'_-\otimes F\big)\delta\boldsymbol{\tau} + \mathcal{M}_-\big(S'_-\mathfrak{F}_{\boldsymbol{\tau}}\otimes U\big),$$

we can conclude that $h^\zeta(S'_-(F\boldsymbol{\tau})) = 0$, by an induction on the size of the graph τ. Hence, $k^\zeta \in G^-_{\mathrm{ad}}$.

– The negative twisted antipode S'_- is defined so as to have identity (7.3) for $x = 0$. Indeed, since $\Pi^\zeta(\bullet^{0,\beta}) \equiv 1$ for all β, one has from the defining relation (7.1) for the twisted antipode, for any $\boldsymbol{\tau} \in \mathcal{U}_{<0}$,

$$\begin{aligned}\mathbb{E}\big[\big({}^{k^\zeta}\Pi^\zeta\boldsymbol{\tau}\big)(0)\big] &= \sum_{\boldsymbol{\varphi}\trianglelefteq\boldsymbol{\tau}} h^\zeta\big(S'_-(\boldsymbol{\varphi})\big)\,\mathbb{E}\big[\big(\Pi^\zeta(\boldsymbol{\tau}/^-\boldsymbol{\varphi})\big)(0)\big] \\ &= \sum_{\boldsymbol{\varphi}\trianglelefteq\boldsymbol{\tau}} h^\zeta\big(S'_-(\boldsymbol{\varphi})\big)\,h^\zeta(\boldsymbol{\tau}/^-\boldsymbol{\varphi}) \\ &= h^\zeta\big(\mathcal{M}(S'_-\otimes \mathrm{Id})\delta\boldsymbol{\tau}\big) = h^\zeta\big(S'_-\tau\big)\big(h^\zeta(\bullet^{0,|\boldsymbol{\tau}|}) - 1\big) = 0.\end{aligned}$$

Recall that the homogeneity and grading notions on U and T are different. It is the homogeneity of $\boldsymbol{\tau}$, seen as an element of T, that appears in $\bullet^{0,|\boldsymbol{\tau}|}$. It is elementary to go from $\mathbb{E}[({}^{k^\zeta}\Pi^\zeta\tau)(0)] = 0$ to $\mathbb{E}[({}^{k^\zeta}\Pi^\zeta\tau)(x)] = 0$ for all $x \in \mathbb{R}\times\mathbb{R}^d$, using the probabilistic translation invariance property of Π^ζ. ■

Remark. Note that the co-interaction identity between δ and δ^- implies that we have

$$ {}^{k_1\star k_2}\Pi^\zeta = {}^{k_1}\big({}^{k_2}\Pi^\zeta\big) \tag{7.4}$$

for any two characters k_1, k_2 on U^-. There is no other character k on U^- than $h^\zeta\circ S'_-$ such that the renormalized naive interpretation operator ${}^k\Pi^\zeta := (k\otimes\Pi^\zeta)\delta$ has property (7.3) of Theorem 44. The uniqueness claim amounts to proving that for any non-null character $k \neq 1$ there exists an element $\boldsymbol{\tau}\in U$ such that $\mathbb{E}[({}^{k\star k^\zeta}\Pi^\zeta\boldsymbol{\tau})(0)] \neq 0$. See the second part of the proof of [20, Theorem 6.18].

Assume now that $\zeta = \xi_\varepsilon$ is the regularized version of a random irregular noise ξ, centered and translation invariant, and write Π^ε for Π^{ξ_ε}. The BHZ character h from (7.2) becomes ε-dependent as well. Set

$$k_\varepsilon := h^{\xi_\varepsilon}\circ S'_-.$$

Identity (7.4) tells us that if the maps ${}^{k_\varepsilon}\Pi^\varepsilon$ converge to a limit when ε goes to zero, then, for any character k on U^-, the renormalized interpretation map ${}^{k\star k_\varepsilon}\Pi^\varepsilon$ is also converging. There is thus a whole class of converging renormalization schemes indexed by the group of characters of U^-, if there is a single converging renormalization scheme. If we insist on building **K**-admissible models, this provides a family of convergent models indexed by the renormalization group G^-_{ad}.

Recall the arguments in Section 6. We say that the family of smooth cylindrical functions $F = \{F(\mathfrak{b}^*)\}_{\mathfrak{b}\in\mathrm{N}^0}$ is a nonlinearity. Denote by $u = \textsc{Sol}(\xi; F)$ for the solution to the PDE

$$(\partial_{x_0} - \Delta_{x'} + 1)u = \sum_{\mathfrak{b}\in\mathrm{N}^0} F(\mathfrak{b}^*)(u, \partial_{x'}u)\Pi^\xi\mathfrak{b}$$

driven by the smooth noise ξ_ε, associated with a given initial condition. The arguments in Section 6.3 imply that the nonlinearity $F = \{F(\mathfrak{b}^*)\}_{\mathfrak{b}\in\mathrm{N}^0}$ can be extended to smooth cylindrical functions $F(\boldsymbol{\tau}^*)$ for any $\boldsymbol{\tau} \in \mathcal{V}$ by (6.25), and the group G_{ad}^- acts on the set $\mathfrak{F}$ of nonlinearities by $F^{(k)} := F \circ \tilde{k}^*$ as in Corollary 41. The renormalization group acquires a dynamical meaning from Theorem 43 if one notices that

$$\big\{\mathrm{SOL}\big(\xi_\varepsilon; F^{(k)}\big)\big\}_{k\in G_{\mathrm{ad}}^-} = \big\{\mathrm{SOL}\big(\xi_\varepsilon; F^{(k_\varepsilon * k)}\big)\big\}_{k\in G^-} = \big\{\mathrm{SOL}\big(\xi_\varepsilon; \big(F^{(k_\varepsilon)}\big)^{(k)}\big)\big\}_{k\in G^-}$$

for any fixed positive ε. This remark tells us that the *family* of solutions of the singular stochastic PDE (1.4) is parametrized by the subset $(F^{(k)})_{k\in G^-}$ of the space $\mathfrak{F}$ of nonlinearities. This remains true at the limit when $\varepsilon > 0$ goes to 0. We will see in Section 8 that this subset is actually a finite-dimensional immersed manifold.

8. The manifold of solutions

We take for granted in this section the convergence result of Chandra and Hairer from [26] and work with the limit random admissible model $\mathsf{M} = (\mathrm{g}, \Pi)$, obtained as a limit in probability of the renormalized naive models ${}^{k_\varepsilon}\mathsf{M}^\varepsilon$ when $\varepsilon > 0$ goes to 0. Recall from equality (6.26) the expression of $\mathscr{F}^{(k)}(\boldsymbol{u})$ for $k \in G_{\mathrm{ad}}^-$. Pick $\eta \in (0, \beta_0 + 2]$, $\gamma \in (-\beta_0, 2)$ and an initial condition $u_0 \in \mathcal{C}^\eta(\mathbb{T}^d)$. Write $\boldsymbol{u}(k) \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g})$ for the solution to the equation

$$\boldsymbol{u}(k) = P_\gamma u_0 + \mathcal{P}^{\mathsf{M}}_{t_0}\big(\mathscr{F}^{(k)}(\boldsymbol{u}(k))\big) =: \Psi_k\big(\boldsymbol{u}(k)\big),$$

and set

$$u(k) := \mathbf{R}^{\mathsf{M}} C_{t_0}\big(\boldsymbol{u}(k)\big).$$

By continuity of the solution map the family of functions $\{u(k)\}_{k\in G_{\mathrm{ad}}^-}$ coincides with the limit of the family

$$\big\{\mathrm{SOL}\big(\xi_\varepsilon; F^{(k * k_\varepsilon)}\big)\big\}_{k\in G_{\mathrm{ad}}^-}.$$

Note that Ψ_k depends linearly, hence smoothly, on k. We saw in Theorem 23 in Section 4 that, given a bounded set of nonlinearities in C^∞, there exists a positive time horizon t_0 such that the "integral" map Ψ_k is a contraction from $\mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g})$, uniformly with respect to the nonlinearities in the given bounded set. So, the continuous linear map $(\mathrm{Id} - \partial_{\boldsymbol{u}}\Psi_k)$, from the Banach space $\mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g})$ into itself, has a continuous inverse, given under the form of the classical Neumann series. The map $(\mathrm{Id} - \partial_{\boldsymbol{u}}\Psi_k)$ is thus a continuous isomorphism of $\mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T, \mathrm{g})$ by the open mapping theorem. It is then a direct consequence of the implicit function theorem that the unique fixed point $\boldsymbol{u}(k)$ of the equation

$$\boldsymbol{u}(k) = \Psi_k\big(\boldsymbol{u}(k)\big)$$

is a smooth function of $k \in G_{\mathrm{ad}}^-$.

Proposition 45. *The family $\{u(k)\}_{k\in G_{\mathrm{ad}}^-}$ forms a finite-dimensional immersed submanifold of $\mathcal{C}^\eta(\mathbb{R} \times \mathbb{R}^d)$, where η is the parameter chosen in Theorem 43.*

Proof. It suffices from the implicit function theorem to see that $D_k\boldsymbol{u}(k)$, the derivative of $\boldsymbol{u}(k)$ with respect to k, has constant rank; this follows from the linearity of the reconstruction map if we can see that $D_k\boldsymbol{u}(k)$ is injective. (The reconstruction map is not injective without further assumptions.) The linear map $D_k\boldsymbol{u}(k)$ sends the tangent space to G^-_{ad} at the identity into $\mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T,\mathrm{g})$. But picking h in that tangent space and setting $\boldsymbol{v} := (D_k\boldsymbol{u}(k))(h) \in \mathcal{D}^{\gamma,\eta}_{(0,t_0)}(T,\mathrm{g})$, the modelled distribution $\boldsymbol{v}$ cannot be null unless $h=0$, since $\boldsymbol{v}$ is the solution to the affine equation

$$\boldsymbol{v} = \mathcal{P}^{\mathsf{M}}_{t_0}\Bigg(D_k(\mathcal{F}^{(k)})(\boldsymbol{u}(k))\boldsymbol{v} + \sum_{\boldsymbol{\tau}\in\mathcal{B},\,|\boldsymbol{\tau}|<0} \frac{h(\boldsymbol{\tau})}{S(\boldsymbol{\tau})} F(\boldsymbol{\tau}^*)(\boldsymbol{u}(k), D\boldsymbol{u}(k))\Bigg). \qquad \blacksquare$$

Remark. The use of the implicit function theorem actually shows that the solution $\boldsymbol{u}$ of the equation

$$\boldsymbol{u} = P_\gamma u_0 + \mathcal{P}^{\mathsf{M}}_{t_0}\big(f^\star(\boldsymbol{u})\Xi + g^\star\big(\boldsymbol{u}, D\boldsymbol{u}\big)\big) \tag{8.1}$$

is a smooth function of $f, g \in C^n$ for n large enough. This gives a direct access to Taylor expansions in small noise, where f is replaced by af, for a small positive parameter a, or if f is the value at $a=0$ of a smooth family $f(a,\cdot) \in C^\infty$, as the solution $\boldsymbol{u}$ happens then to be a smooth function of the expansion parameter a. Elementary classical calculus is used to see that the derivatives of $\boldsymbol{u}$ with respect to the parameter a are solutions of affine equations obtained by formal differentiation of equation (8.1) with respect to the parameter. This kind of questions has a long history, under the name "stochastic Taylor expansion" in a stochastic calculus setting – after seminal works by Azencott [1] and Ben Arous [14], where it was used together with the stationary phase method on Wiener space to get heat kernel estimates for elliptic and sub-elliptic diffusions. Inahama and Kawabi extended the approach to a rough paths setting in [66], and Friz, Gassiat, and Pigato made a first use of this type of ideas in a regularity structures setting in [42]. The result of Proposition 45 holds for all subcritical singular stochastic PDEs, with the above straightforward proof.

9. Building regularity and renormalization structures

In the end, for the above results to hold, we require from the regularity structure $\mathcal{T}$ and the renormalization structure $\mathcal{U}$ that they satisfy the different assumptions introduced along the way for different purposes. We summarize them in Table 1, with a quick description of what they are useful for.

Following Bruned, Hairer, and Zambotti [20], we describe in this section a setting tailor-made for the study of the generalized KPZ equation (6.1) where all these conditions hold true. We introduce a homogeneity map on the decorated trees from Section 6.1.

Definition. Let $\mathfrak{T}_{\mathrm{n}}$ and $\mathfrak{T}_{\mathrm{e}}$ be abstract finite sets, equipped with homogeneity maps $|\cdot| : \mathfrak{T}_{\mathrm{n}}, \mathfrak{T}_{\mathrm{e}} \to \mathbb{R}$.

Assumption	Section	What it is useful for
A1–A2	2.3	Inclusion of the polynomial structure in our regularity structures.
A3	2.4	Product between T_X and T.
B1	3.2	Actions of $\Delta^{(+)}$ on $\mathcal{I}_n^{(+)}$.
B2	3.5	Induction structure on Δ for building admissible models.
C	5.2	Compatibility between the maps $\delta^{(+)}$ and $\mathcal{I}_n^{(+)}$.
D1–D2	6.1, 6.2	Largeness of the basis $\mathcal{B}$ of T and U.
D3	6.3	Compatibility between multi-pre-Lie and renormalization structures.
E	7	Structure assumption on U^- and induction structure on δ.

Table 1. List of assumptions.

- On the sets $\mathsf{N} := \mathfrak{T}_{\mathrm{n}} \times \mathbb{N} \times \mathbb{N}^d$ and $\mathsf{E} := \mathfrak{T}_{\mathrm{e}} \times \mathbb{N} \times \mathbb{N}^d$, the homogeneity maps are extended by
$$\begin{cases} |k|_n := |n| + |k|_{\mathfrak{s}}, & (n,k) \in \mathsf{N}, \\ |\ell|_e := |e| - |\ell|_{\mathfrak{s}}, & (e,\ell) \in \mathsf{E}. \end{cases}$$
- The *naive homogeneity* of a decorated tree $\tau_e^{\mathfrak{n}}$ is defined by
$$|\tau_e^{\mathfrak{n}}|' := \sum_{e \in E_\tau} |\mathfrak{e}(e)|_{\mathfrak{t}_{\mathrm{e}}(e)} + \sum_{n \in N_\tau} |\mathfrak{n}(n)|_{\mathfrak{t}_{\mathrm{n}}(n)}.$$

We start from the sets
$$\mathfrak{T}_{\mathrm{n}} = \{\bullet, \circ\}, \quad \mathfrak{T}_{\mathrm{e}} = \{\mathcal{I}\}$$
for an abstract symbol $\mathcal{I}$ – we use on purpose the same symbol as the abstract integration map from Section 3.2. The node-type set $\mathfrak{T}_{\mathrm{n}}$ is enlarged later. The two elements $\bullet$ and $\circ$ of $\mathfrak{T}_{\mathrm{n}}$ represent the monomial $\mathbf{1} = X^0$ and the noise Ξ, respectively. The set $\mathfrak{T}_{\mathrm{e}}$ consists of only one integration operator $\mathcal{I}$. Each element has homogeneity
$$|\bullet| = 0, \quad |\circ| = \beta_0, \quad |\mathcal{I}| = 2,$$
where $\beta_0 \in (-2,0)$ is the regularity of the noise ζ in the equation. (Would the equation under study involve several noises with different regularities, we would introduce several $\circ$ symbols with the corresponding homogeneities.) Given that the polynomial structure is needed to encode at a regularity structure level the term $P_\gamma u_0$ describing the propagation of the initial condition, and the piece of $\mathcal{P}_t^{\mathsf{M}}$ taking values in the polynomial regularity structure, the use of trees with a node decoration encoding multiplication by polynomials appears as natural. On the other hand, the use of edge decorations for equations that do not involve derivatives of the solution in their formulation, like the generalized (PAM) equation
$$(\partial_t - \Delta_x)u = f(u)\zeta,$$

may look strange. The necessity to use edge decorations to encode derivatives of quantities of the form $\mathcal{I}(\cdot)$, even in such a case, comes from the renormalization process implemented in this setting, as the latter involves Taylor expansions.

As said in Section 6, the final form of a generic element of our regularity structures will be the datum of a decorated tree together with a colouring and an additional decoration

$$\mathfrak{o} : N_\tau \to \mathbb{Z}[\beta_0],$$

which plays an important role in the compatibility condition between regularity and renormalization structures from Definition 27. In a nutshell, this additional decoration will keep track of the naive homogeneity of the "diverging" trees that will be extracted by the renormalization map δ. This is what will allow to have maps $\delta^{(+)}$ satisfying the fundamental conditions

$$\delta^{(+)} T_\beta^{(+)} \subset U^- \otimes T_\beta^{(+)} \quad (\beta \in A^{(+)})$$

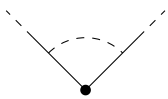
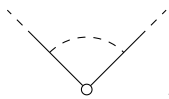
involved in the definition of compatible regularity and renormalization structures. So, one should not be surprised that we will use the naive homogeneity to define the gradings in U and U^- and a different notion of homogeneity in T and T^+, taking into account the $\mathfrak{o}$-decorations. The discussion will be general enough for the reader to see what needs to be added to deal with the general case.

9.1. Rules and extended decoration

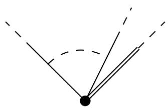
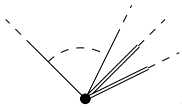
Working with the set of all decorated trees as a candidate for a regularity structure is not reasonable, and we first identify a few notions that help clarifying the matter. Recall the abstract self-explaining formulation

$$\boldsymbol{v} = \mathcal{I}\big(f^\star(\boldsymbol{v})\Xi + g^\star(\boldsymbol{v}, \partial\boldsymbol{v})\big) + (T_X) \tag{9.1}$$

of the generalized (KPZ) equation. In the present tree, setting the $\star$ product is given by the "*joining*" operator $\mathcal{J}$ on trees. If one wants to make sense of Picard iteration within the concrete regularity structure, one needs to make sense of a number of recursive relations – recall the subcomodules introduced in Section 4.4 and see the pictures in Section 6.1. General constraints of this type come under the name of *rule*, that is, the definition for each node type $\mathfrak{b} \in \mathfrak{T}_{\mathrm{n}}$ of constraints on which kind of tuples of edges $\{e^i = (e^i_-, e^i_+)\}_i$ can have $\mathfrak{t}_{\mathrm{n}}(e^i_-) = \mathfrak{b}$, for all i, in a tree allowed by the rule. The choice of a rule is determined by the equation under consideration. Consider the right-hand side of equation (9.1). Making sense of the nonlinear term $f^\star(\boldsymbol{v})\Xi + g^\star_0(\boldsymbol{v})$ requires that one can find $\mathcal{J}(\mathcal{I}(\cdot), \dots, \mathcal{I}(\cdot))X^n$ or $\mathcal{J}(\mathcal{I}(\cdot), \dots, \mathcal{I}(\cdot))X^n\,\Xi$, within the trees allowed by the rule; that is, the corresponding nodes are of the form

 or  .

Making sense of the other terms $g_2^\star(\boldsymbol{v}) \star (D\boldsymbol{v})^{\star 2} + g_1^\star(\boldsymbol{v}) \star (D\boldsymbol{v})$ requires that one can find $\mathcal{J}(\mathcal{I}(\cdot), \dots, \mathcal{I}(\cdot), \mathcal{I}_{e_i}(\cdot))X^n$, or $\mathcal{J}(\mathcal{I}(\cdot), \dots, \mathcal{I}(\cdot), \mathcal{I}_{e_i}(\cdot), \mathcal{I}_{e_j}(\cdot))X^n$ for some $i, j = 1, \dots, d$ within the trees allowed by the rule, so each node of the corresponding elements of C has the form

 or  .

The operators $\mathcal{I}_{e_i}$ are represented by the double line in the above picture. Given a rule, a decorated *conforming tree* is a tree such that all nodes of the tree, except perhaps the root, satisfy the rule. Denote by

$$C$$

the set of conforming trees. If all nodes of the tree satisfy the rule, the tree is called *strongly conforming*. We denote by

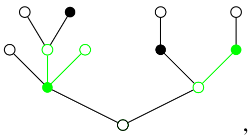
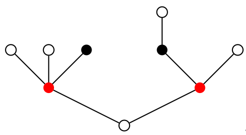
$$SC$$

the set of strongly conforming trees. A rule is said to be *normal* if any subtree of a strongly conforming tree is also strongly conforming.

To construct regularity and renormalization structures, the rooted decorated trees obtained from the above iterations are not sufficient. Another important operation is the *contraction* of rooted trees, involved in the definition of the splitting maps Δ and δ. Given a typed rooted tree τ and a family φ of disjoint typed subtrees of τ, we use the notation

$$\tau/^{\mathrm{red}}\varphi$$

to denote the typed rooted tree obtained by identifying each subtree τ_i with a single node $\color{red}\bullet$ with red colour in the quotient tree. Here is an example, with φ in green,

$\tau =$  , $\tau/^{\mathrm{red}}\varphi =$  .

We allow such an operation for the set SC of strongly conforming trees. Precisely, if each connected component of φ belongs to SC, then we assume that $\tau/^{\mathrm{red}}\varphi \in SC$. Hence, each element of SC is a rooted decorated tree with a node-type set

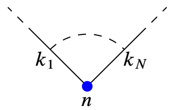
$$\mathfrak{T}_{\mathrm{n}}^{SC} = \{\bullet, {\color{red}\bullet}, \circ\}.$$

The analytic role of $\color{red}\bullet$ is the same as that of $\bullet$. In particular, the homogeneity of $\color{red}\bullet$ is 0. This is an example of the colouring of the tree. Only decorated trees without red colour appear in the analysis of the well-posedness problem (9.1), but colours are used in the definition of the splitting maps in the renormalization structure.

Recall from Assumption B1 and Section 3.5 that the algebra T^+ is spanned by elements of the form

$$X^n \prod_{i=1}^{N} \mathcal{I}^+_{k_i}(\tau_i), \tag{9.2}$$

where $n \in \mathbb{N} \times \mathbb{N}^d$, $k_1, \dots, k_N \in \mathbb{N} \times \mathbb{N}^d$, and $\tau_1, \dots, \tau_N \in SC$. It is convenient to consider an element like (9.2) as a tree by interpreting $\mathcal{I}^+_k$ as the planting operator like $\mathcal{I}_k$ and the product $\prod$ as the tree product $\mathcal{J}$. To distinguish such trees from elements of SC, we give a blue colour to their roots, encoding in this way the $+$ sign in $\mathcal{I}^+$.



The set C consists of such trees, where we see that the rule is broken at the root. This is because C is only conforming, not strongly conforming. Each element of C is thus a rooted decorated tree with a node-type set

$$\mathfrak{T}^C_{\mathfrak{n}} = \{\bullet, {\color{red}\bullet}, {\color{blue}\bullet}, \circ\}.$$

A node of a conforming tree has the type ${\color{blue}\bullet}$ if and only if it is a root. The homogeneity of ${\color{blue}\bullet}$ is 0. The trees with a blue root will only be involved in the description of the space T^+.

A rule is said to be *subcritical* if for any $\gamma \in \mathbb{R}$, only finitely many elements of SC have naive homogeneity less than γ. A *complete* rule will guarantee that a rooted decorated tree obtained from the contraction of a strongly conforming tree by extracting "diverging" pieces, and changing the decorations accordingly, will still be strongly conforming. Reference [20, Proposition 5.21] ensures that *any normal subcritical rule can be extended into a normal subcritical complete rule*. We take this result for granted and do not reprove it here. The above rule on the set of decorated trees is normal, subcritical, and complete.

To construct compatible regularity and a renormalization structure, we introduce an additional decoration. Denote by N^{red}_τ the subset of N_τ consisting of the nodes with type ${\color{red}\bullet}$.

Definition. A tree with *extended decoration* is a rooted decorated tree $\tau^{\mathfrak{n}}_{e}$ with a map

$$\mathfrak{o} : N^{\text{red}}_\tau \to \mathbb{Z}[\beta_0].$$

We write $\boldsymbol{\tau} = \tau^{\mathfrak{n},\mathfrak{o}}_{e}$ for a generic tree with extended decoration. The *extended homogeneity* of such a tree is defined by

$$\left|\tau^{\mathfrak{n},\mathfrak{o}}_{e}\right| := \sum_{e \in E_\tau} \left|\mathfrak{e}(e)\right|_{\mathfrak{t}_e(e)} + \sum_{n \in N_\tau} \left|\mathfrak{n}(n)\right|_{\mathfrak{t}_{\mathfrak{n}}(n)} + \sum_{n \in N^{\text{red}}_\tau} \mathfrak{o}(n).$$

We extend the naive homogeneity to the set of decorated trees with an extended decoration setting

$$\left|\tau^{\mathfrak{n},\mathfrak{o}}_{e}\right|' := \left|\tau^{\mathfrak{n}}_{e}\right|'.$$

Note here that only trees without $\mathfrak{o}$-decoration, that is, $\mathfrak{o} = 0$, appear in the analysis of the fixed-point problem (9.1). Indeed, the trees without $\mathfrak{o}$-decoration are stable under the coproducts (Δ^+, Δ) defined below. The $\mathfrak{o}$-decoration is only involved in the analysis of the renormalization procedure and the associated convergence problem – see the second equality of (9.5). Without it, the condition (5.3) for the compatibility of $\mathcal{T}$ and $\mathcal{U}$ does not hold. We define the set

$$\mathcal{SC}$$

of strongly conforming trees with $\mathfrak{o}$-decoration as the minimal set which contains SC and such that the vector space spanned by $\mathcal{SC}$ is stable under all the coproducts defined below. (One could also consider $\mathcal{SC}$ as a set of rooted decorated trees with node-type set

$$\mathfrak{T}_{\mathrm{n}}^{\mathcal{SC}} = \{\bullet, \circ\} \cup \{{\color{red}\bullet}^{\cdot,\alpha}\}_{\alpha\in\mathbb{Z}[\beta_0]}.$$

We used such an identification in Section 6. In the present section, we treat $\mathfrak{o}$ as a decoration, rather than as part of a node type.) Similarly, we define

$$\mathcal{C}$$

as the set of decorated trees with extended decorations of the form (9.2), where $\tau_1, \dots, \tau_N \in \mathcal{SC}$. We use the bold symbol $\boldsymbol{\tau}$ to denote a generic element of $\mathcal{SC}$ or $\mathcal{C}$. The above rule on the set of extended decorated trees is normal, subcritical, and complete. (The subcriticality of the rule on this set of trees with extended decorations comes from the fact that, for any fixed $\gamma \in \mathbb{R}$, the decoration α of trees with extended homogeneity less than γ, will only range in the set of homogeneities of subtrees of strongly conforming trees $\tau_e^{\mathfrak{n}}$ with homogeneity less than γ.)

9.2. Coproducts

We define coproducts in the spaces of rooted decorated trees. This requires first that we define what we mean by "subtrees" and "subforests". Recall that the type sets

$$\mathfrak{T}_{\mathrm{n}}^{SC} = \{\bullet, {\color{red}\bullet}, \circ\}, \quad \mathfrak{T}_{\mathrm{n}}^{C} = \{\bullet, {\color{red}\bullet}, {\color{blue}\bullet}, \circ\}, \quad \mathfrak{T}_{\mathrm{e}} = \{\mathcal{I}\}$$

are fixed. Given a typed rooted tree τ, a nonempty connected subgraph of τ is called a *subtree* if it inherits from τ its type map. Any possibly empty family of disjoint subtrees of τ is called a *subforest*. Given a rooted tree τ and a subforest $\varphi = \{\tau_1, \dots, \tau_m\}$, we use the notation

$$\tau/\varphi$$

to denote the rooted tree obtained by identifying each subtree τ_i with a single node with node type $\bullet$ in the quotient tree. Precisely, writing $y \sim_\varphi z$ if y and z are in the same connected component of φ, we define τ/φ as the tree consisting of the node set $N_\tau/\sim_\varphi$ and the edge set $E_\tau \setminus E_\varphi$. Moreover, we write

$$\tau/^{\mathrm{red}}\varphi, \quad \text{or} \quad \tau/^{\mathrm{blue}}\varphi$$

if we give a corresponding colour to the nodes of φ in the quotient tree.

- For any function $f : N_\tau \to \mathbb{N} \times \mathbb{N}^d$, define the function $[f]_\varphi$ on $N_{\tau/\varphi}$ by

$$[f]_\varphi\big([x]\big) := \sum_{y \sim_\varphi x} f(y),$$

where $[x]$ denotes the equivalence class of $x \in N_\tau$.

- Denote by $\partial\varphi$ the leaves of φ, that is, the set of edges $(x, y) \in E_\tau$ such that $x \in N_\varphi$ and $y \in N_\tau \setminus N_\varphi$. For any function $g : \partial\varphi \to \mathbb{N} \times \mathbb{N}^d$, define the function πg on N_τ by setting

$$(\pi g)(x) := \sum_{e=(x,y)\in\partial\varphi} g(e).$$

- For any decorations $\mathfrak{n}_\varphi$ and $\mathfrak{e}_\varphi$ on φ, define the function $\mathfrak{o}(\varphi, \mathfrak{n}_\varphi, \mathfrak{e}_\varphi) : N_{\tau/\varphi} \to \mathbb{Z}[\beta_0]$ by

$$\mathfrak{o}\big(\varphi, \mathfrak{n}_\varphi, \mathfrak{e}_\varphi\big)\big([\tau_j]\big) = \Big|(\tau_j)^{\mathfrak{n}_\varphi|_{\tau_j}}_{\mathfrak{e}_\varphi|_{\tau_j}}\Big|'$$

for each $1 \le j \le m$, and $\mathfrak{o}(\varphi, \mathfrak{n}_\varphi, \mathfrak{e}_\varphi) = 0$ outside $[\varphi]$.

Define

$$T := \mathrm{span}(\mathcal{SC}), \quad \mathsf{T}^+ := \mathrm{span}(\mathcal{C}), \quad \mathsf{U}^- := \mathbb{R}[\mathcal{SC}].$$

Note that T^+ is an algebra with the tree product and unit $\mathbf{1}_+ := \bullet^0$, and U^- is an algebra with the forest product and unit $\mathbf{1}_- := \emptyset$. The space T^+ will be built from the side space T^+ and the space U^- from the side space U^-. Similarly, the different splitting maps defining a regularity structure and a renormalization structure are built from splitting maps taking values in, or defined on, the spaces T, T^+, U^-.

Definition. We introduce three splitting operators.

(1) The linear map

$$\mathsf{D} : T \to T \otimes \mathsf{T}^+$$

is defined for $\tau^{\mathfrak{n},\mathfrak{o}}_{\mathfrak{e}} \in \mathcal{SC}$ by

$$\mathsf{D}\tau^{\mathfrak{n},\mathfrak{o}}_{\mathfrak{e}} := \sum_{\mu} \sum_{\mathfrak{n}_\mu, \mathfrak{e}'_{\partial\mu}} \frac{1}{\mathfrak{e}'_{\partial\mu}!} \binom{\mathfrak{n}}{\mathfrak{n}_\mu} \mu_{\mathfrak{e}}^{\mathfrak{n}_\mu + \pi\mathfrak{e}'_{\partial\mu}, \mathfrak{o}|_\mu} \otimes \big(\tau/^{\mathrm{blue}}\mu\big)^{[\mathfrak{n}-\mathfrak{n}_\mu]_\mu, \mathfrak{o}|_{\tau\setminus\mu}}_{\mathfrak{e}+\mathfrak{e}'_{\partial\mu}}, \tag{9.3}$$

where the first sum is over all subtrees μ of τ which contains the root of τ, and the second sum is over functions $\mathfrak{n} : N_\mu \to \mathbb{N} \times \mathbb{N}^d$, with $\mathfrak{n}_\mu \le \mathfrak{n}$ and functions $\mathfrak{e}'_{\partial\mu} : \partial\mu \to \mathbb{N} \times \mathbb{N}^d$. The algebra morphism

$$\mathsf{D}^+ : \mathsf{T}^+ \to \mathsf{T}^+ \otimes \mathsf{T}^+$$

is defined by the same formula (9.3) for $\tau^{\mathfrak{n},\mathfrak{o}}_{\mathfrak{e}} \in \mathcal{C}$.

(2) The algebra morphism

$$\mathsf{D}^- : \mathsf{U}^- \to \mathsf{U}^- \otimes \mathsf{U}^-$$

is defined by $\mathsf{D}^-(\mathbf{1}_-) := \mathbf{1}_- \otimes \mathbf{1}_-$, and for $\tau_e^{\mathfrak{n},\mathfrak{o}} \in \mathcal{SC}$

$$\mathsf{D}^-\big(\tau_e^{\mathfrak{n},\mathfrak{o}}\big) := \sum_{\varphi} \sum_{\mathfrak{n}_\varphi, \mathfrak{e}'_{\partial\varphi}} \frac{1}{\mathfrak{e}'_{\partial\varphi}!} \binom{\mathfrak{n}}{\mathfrak{n}_\varphi} \varphi_{\mathfrak{e}}^{\mathfrak{n}_\varphi + \pi \mathfrak{e}'_{\partial\varphi}, \mathfrak{o}|_\varphi} \otimes \big(\tau/^{\mathrm{red}}\varphi\big)_{\mathfrak{e}+\mathfrak{e}'_{\partial\varphi}}^{[\mathfrak{n}-\mathfrak{n}]_\varphi, [\mathfrak{o}]_\varphi + \mathfrak{o}(\varphi, \mathfrak{n}_\varphi + \pi \mathfrak{e}'_{\partial\varphi}, \mathfrak{e})},$$

where the first sum is over all subforests φ of τ which contains all red nodes of τ, and the sum over $\mathfrak{n}_\varphi$ and $\mathfrak{e}'_{\partial\varphi}$ is taken as in item 1 of the present definition.

(3) The algebra morphism

$$\overline{\mathsf{D}}^- : \mathsf{T}^+ \to \mathsf{U}^- \otimes \mathsf{T}^+$$

is defined by the same formula as D^-, with the first sum restricted to subforests φ which are disjoint from the root of τ.

Remark. Since the right-hand side of (9.3) may become an infinite series, we have to consider the "bigraded spaces" of rooted decorated trees defined in [20, Section 2.3]. For any collection of vector spaces $\{V_n\}_{n\in\mathbb{N}^2}$, we denote by $V = \boxplus_{n\in\mathbb{N}^2} V_n$ the space of all sequences $(v_n)_{n=(n_1,n_2)\in\mathbb{N}^2}$ with $v_n \in V_n$ such that there exist $k \in \mathbb{N}$ and $v_n = 0$ unless $n_2 \le k$. The tensor product of two bigraded spaces $V = \boxplus_{n\in\mathbb{N}^2} V_n$ and $W = \boxplus_{n\in\mathbb{N}^2} W_n$ is defined by

$$V \,\hat{\otimes}\, W := \boxplus_{n\in\mathbb{N}^2} \Big(\bigoplus_{k+\ell=n} V_k \otimes W_\ell \Big).$$

For example, the bigraded space $\mathsf{V} = \boxplus_{n\in\mathbb{N}^2} \mathsf{V}_n$ of rooted decorated trees is given by setting $\mathsf{V}_{(n_1,n_2)}$ as the vector space spanned by all decorated trees $\tau_e^{\mathfrak{n},\mathfrak{o}}$ such that

$$\sum_{e\in E_\tau} |\mathfrak{e}(e)|_{\mathfrak{s}} = n_1 \quad \text{and} \quad |N_\tau \setminus \mathfrak{t}^{-1}\{\color{red}\bullet\color{black}, \color{blue}\bullet\color{black}\}| = n_2.$$

The spaces T, T^+, and U^- above are defined as sub-bigraded spaces of V, and their tensor products are also defined as bigraded spaces in the above sense. As in [20, Lemma 2.14], triangular maps between bigraded spaces are well defined. For any bigraded spaces V and W, the family of linear maps $A_{mn} : V_n \to W_m$ is called *triangular* if $A_{mn} = 0$ unless $m_1 \ge n_1$ and $m_2 \le n_2$. The the linear map from V to W

$$A\big((v_n)_{n\in\mathbb{N}^2}\big) := \Big(\sum_{n\in\mathbb{N}^2} A_{mn} v_n\Big)_{m\in\mathbb{N}^2}$$

is well defined. The maps D, D^+, D^-, and $\overline{\mathsf{D}}^-$ above are well defined as triangular maps. In the following, when dealing with infinite series, we use these facts implicitly. In Section 9.3, we introduce some truncation maps which reduce infinite series to finite sums.

As suggested by the target spaces of the preceding maps, the splitting map Δ will be constructed from D and the map Δ^+ from D^+, the maps δ and δ^- from D^-, and the map δ^+ from $\overline{\mathsf{D}}^-$. Only trees with blue roots appear on the right-hand side of the tensor products defining D. This is consistent with the fact that the trees with blue roots will represent later elements of T^+. The restriction on the choice of φ to subforests which are disjoint from the root in the definition of $\overline{\mathsf{D}}^-$ ensures that it takes values in $\mathsf{U}^- \otimes \mathsf{T}^+$ and that the multiplicative property

$$\overline{\mathsf{D}}^-(\boldsymbol{\tau}\boldsymbol{\sigma}) = \big(\overline{\mathsf{D}}^-\boldsymbol{\tau}\big)\big(\overline{\mathsf{D}}^-\boldsymbol{\sigma}\big)$$

holds. This reflects the fact that the product of two functions

$$\mathsf{g}(\boldsymbol{\tau})\,\mathsf{g}(\boldsymbol{\sigma}), \quad \boldsymbol{\tau}, \boldsymbol{\sigma} \in \mathscr{C}$$

does not cause any new renormalization.

Remark. Keep in mind that the elements of U^- are meant to be evaluated by characters of U^-, and turned to numbers, while elements of U are meant to be turned to distributions. This is done jointly in a renormalized naive model $(k \otimes \Pi^\zeta)\delta$. Recall that the problem of renormalization comes from the fact that the kernel of the operator $\mathbf{K}$ explodes on the diagonal. The building block of the renormalization operations δ and δ^- is best understood in the light of the following archetype problem. Let $g : ([0,1]^d)^n \to \mathbb{R}$ be a function that is smooth outside the deep diagonal $\mathrm{diag} := \{\mathbf{z} = (z_1, \dots, z_n) \in ([0,1]^d)^n; z_1 = \cdots = z_n\}$, near which it behaves as $|\mathbf{z} - (z_1, \dots, z_1)|^{-a}$, for an exponent $a > d$. The function g is not integrable in any neighbourhood of the deep diagonal, so it only makes sense as a distribution on $([0,1]^d)^n \setminus \mathrm{diag}$

$$\int_{([0,1]^d)^n} g(\mathbf{z}) f(\mathbf{z}) d\mathbf{z},$$

for f smooth, with support with empty intersection with the deep diagonal. Can we define a distribution Λ on $([0,1]^d)^n$ that extends this distribution? This can be done by defining Λ on $([0,1]^d)^n$

$$(\Lambda, \psi) = \int_{([0,1]^d)^n} g(\mathbf{z}) \left(\psi(\mathbf{z}) - \psi(\mathbf{z}_1) - \cdots - \frac{(\mathbf{z} - \mathbf{z}_1)^{[a-d]}}{[a-d]!}\, \psi^{([a-d])}(\mathbf{z}_1) \right) d\mathbf{z},$$

for any smooth function ψ on $([0,1]^d)^n$. This formula defines indeed a distribution, which coincides with the distribution associated with g outside the deep diagonal, since the $\psi^{[\ell]}(\mathbf{z}_1)$ are null for functions with compact support with null intersection with diag. Taylor expansion appears as the building block of this extension procedure. In this parallel, τ has two pieces, g and f, so the role of φ in D^- would be played by either of them, and the role of the projector p_- in δ^-, defined below, would select only the diverging term. The term $\varphi \otimes (\tau/^-\varphi)$ in $\delta^-\tau$ would precisely correspond to a term $g(\mathbf{z})\, \frac{(\mathbf{z}-\mathbf{z}_1)^n}{n!}\, f^{(n)}(\mathbf{z}_1)$ in the integral defining Λ. A formula like the above defining relation for D^- appears if one deals with a multiple integral where several subintegrals define functions of their external variables of the same kind as g, and one uses a similar kind of extension procedure as above.

The following lemma is proved in Appendix C.3.

Lemma 46. *One has the co-associativity formulas*

$$(\mathsf{D}\otimes \mathrm{Id})\mathsf{D} = (\mathrm{Id}\otimes \mathsf{D}^+)\mathsf{D}, \qquad (\mathsf{D}^+\otimes \mathrm{Id})\mathsf{D}^+ = (\mathrm{Id}\otimes \mathsf{D}^+)\mathsf{D}^+,$$
$$(\mathsf{D}^-\otimes \mathrm{Id})\mathsf{D}^- = \big(\mathrm{Id}\otimes \mathsf{D}^-\big)\mathsf{D}^-, \quad (\mathsf{D}^-\otimes \mathrm{Id})\overline{\mathsf{D}}^- = \big(\mathrm{Id}\otimes \overline{\mathsf{D}}^-\big)\overline{\mathsf{D}}^-.$$

Moreover, one has the cointeraction formulas

$$\mathcal{M}^{(13)}\big(\mathsf{D}^-\otimes \overline{\mathsf{D}}^-\big)\mathsf{D} = (\mathrm{Id}\otimes \mathsf{D})\mathsf{D}^-,$$
$$\mathcal{M}^{(13)}\big(\overline{\mathsf{D}}^-\otimes \overline{\mathsf{D}}^-\big)\mathsf{D}^+ = (\mathrm{Id}\otimes \mathsf{D}^+)\overline{\mathsf{D}}^-.$$

9.3. Regularity and renormalization structures

We define the Hopf algebra parts of regularity and renormalization structures, from the side spaces T^+ and U^-. We use the shorthand notation $\mathcal{I}_n^+(\tau)$ to denote the tree $\mathcal{I}_n(\tau)$ with a blue root, with $\mathcal{I}_n(\tau)$ standing for $\mathcal{I}(\tau)$ with decoration n on the edge outgoing from the root. We define subsets $\mathcal{C}^+\subset\mathcal{C}$ and $\mathcal{SC}^-\subset\mathcal{SC}$ by

$$\mathcal{C}^+ := \big\{X^n\mathcal{J}\big(\mathcal{I}_{m_1}^+(\boldsymbol{\tau}_1),\dots,\mathcal{I}_{m_b}^+(\boldsymbol{\tau}_b)\big)\in\mathcal{C};\ \big|\mathcal{I}_{m_j}^+(\boldsymbol{\tau}_j)\big|>0,\ \text{for any } j=1,\dots,b\big\},$$
$$\mathcal{SC}^- := \big\{\boldsymbol{\tau}\in\mathcal{SC};\ |\boldsymbol{\tau}|'<0\big\}$$

and set

$$T=U=\mathrm{span}(\mathcal{SC}),\quad T^+:=\mathrm{span}(\mathcal{C}^+),\quad U^-:=\mathbb{R}[\mathcal{SC}^-].$$

Note the use of the two notions of homogeneity in these definitions, the extended homogeneity $|\cdot|$ for T and T^+, and the naive homogeneity $|\cdot|'$ for U and U^-. Denote by

$$p_+:\mathsf{T}^+\to T^+$$

the canonical projection, and define an algebra morphism

$$p_-:\mathsf{U}^-\to U^-$$

setting

$$p_-(\boldsymbol{\tau}) := \begin{cases} \mathbf{1}_-, & \text{for } \boldsymbol{\tau}=\mathbf{1}_-, \bullet^{0,\alpha},\\ \boldsymbol{\tau}, & \text{for } \boldsymbol{\tau}\in\mathcal{SC}^-,\\ 0, & \text{for } \boldsymbol{\tau}\in\mathcal{SC}\setminus\big\{\mathcal{SC}^-\cup\{\bullet^{0,\alpha}\}_{\alpha\in\mathbb{Z}[\beta_0]}\big\}.\end{cases}$$

Definition. Define the linear maps

$$\Delta := \big(\mathrm{Id}\otimes p_+\big)\mathsf{D}: T\to T\otimes T^+,\quad \Delta^+ := \big(p_+\otimes p_+\big)\mathsf{D}^+|_{T^+}: T^+\to T^+\otimes T^+,$$
$$\delta := \big(p_-\otimes\mathrm{Id}\big)\mathsf{D}^-|_U: U\to U^-\otimes U,\quad \delta^- := \big(p_-\otimes p_-\big)\mathsf{D}^-|_{U^-}: U^-\to U^-\otimes U^-,$$
$$\delta^+ := (p_-\otimes\mathrm{Id})\overline{\mathsf{D}}^-|_{T^+}: T^+\to U^-\otimes T^+.$$

It follows from the multiplicativity of $p_\pm$ that Δ^+ and $\delta^\pm$ *are algebra morphisms*. The assumption $p_-(\bullet^{0,\mathfrak{o}}) = \mathbf{1}_-$ is needed to ensure the formulas

$$\begin{aligned}\delta\boldsymbol{\tau} &= \mathbf{1}_- \otimes \boldsymbol{\tau} + \sum_{|\boldsymbol{\varphi}|<0} \boldsymbol{\varphi} \otimes (\boldsymbol{\tau}/\boldsymbol{\varphi}),\\ \delta^-\boldsymbol{\sigma} &= \mathbf{1}_- \otimes \boldsymbol{\sigma} + \boldsymbol{\sigma} \otimes \mathbf{1}_- + \sum_{|\boldsymbol{\sigma}|<|\boldsymbol{\psi}|<0} \boldsymbol{\psi} \otimes (\boldsymbol{\sigma}/\boldsymbol{\psi})\end{aligned}$$

for $\boldsymbol{\tau} \in \mathcal{SC}$ and $\boldsymbol{\sigma} \in \mathcal{SC}^-$.

Theorem 47. *Set*

$$\begin{aligned}\mathcal{T} &:= \big((T^+, \Delta^+), (T, \Delta)\big),\\ \mathcal{U} &:= \big((U^-, \delta^-), (U, \delta)\big).\end{aligned}$$

(a) *$\mathcal{T}$ is a regularity structure satisfying Assumptions* A *and* B, *with the grading* $|\cdot|$.

(b) *$\mathcal{U}$ is a renormalization structure satisfying Assumption* E, *with the grading* $|\cdot|'$.

(c) *$\mathcal{T}$ and $\mathcal{U}$ are compatible and satisfy Assumption* C.

(d) *Assumption* D, *the compatibility between the splittings* $\Vdash_{\mathsf{e}}|_U$ *and* δ *holds true.*

Proof. Write as shorthand

$$\begin{aligned}\mathsf{D}\boldsymbol{\tau} \quad \text{or} \quad \mathsf{D}^+\boldsymbol{\tau} &= \sum_i \boldsymbol{\sigma}_i \otimes \boldsymbol{\eta}_i,\\ \mathsf{D}^-\boldsymbol{\tau} \quad \text{or} \quad \bar{\mathsf{D}}^-\boldsymbol{\tau} &= \sum_j \boldsymbol{\varphi}_j \otimes \boldsymbol{\psi}_j,\end{aligned}$$

and note that the following stability formulas of the naive and extended homogeneities. One has

$$|\boldsymbol{\tau}| = |\boldsymbol{\sigma}_i| + |\boldsymbol{\eta}_i|, \tag{9.4}$$

$$|\boldsymbol{\tau}|' = |\boldsymbol{\varphi}_j|' + |\boldsymbol{\psi}_j|', \quad |\boldsymbol{\tau}| = |\boldsymbol{\psi}_j| \tag{9.5}$$

for each i and j. Here, we define $|\mathbf{1}_-|' := 0$.

(a) By the first identity of (9.4),

$$(p_+ \otimes p_+)\mathsf{D}^+ p_+ = (p_+ \otimes p_+)\mathsf{D}^+$$

holds on T^+. Then, one has the comodule property of Δ as follows:

$$\begin{aligned}(\Delta \otimes \mathrm{Id})\Delta &= \big(\mathrm{Id} \otimes p_+ \otimes p_+\big)(\mathsf{D} \otimes \mathrm{Id})\mathsf{D}\\ &= \big(\mathrm{Id} \otimes p_+ \otimes p_+\big)(\mathrm{Id} \otimes \mathsf{D}^+)\mathsf{D}\\ &= \big(\mathrm{Id} \otimes p_+ \otimes p_+\big)(\mathrm{Id} \otimes \mathsf{D}^+)(\mathrm{Id} \otimes p_+)\mathsf{D}\\ &= (\mathrm{Id} \otimes \Delta^+)\Delta.\end{aligned}$$

The co-associativity of Δ^+ is obtained similarly. One gets for free the existence of an antipode on T^+ from the fact that T^+ is a connected graded bialgebra – see Proposition 48 in Appendix B.

(b) The comodule properties of δ and δ^- are obtained by a similar way to (a), since

$$(p_- \otimes p_-)\mathsf{D}^- p_- = (p_- \otimes p_-)\mathsf{D}^-$$

holds on U^-, by identity (9.5). By definition, $\mathbf{1}_-$ is the only element in U^- of 0 homogeneity, so U^- is a connected graded bialgebra.

(c) We prove the cointeraction property

$$\mathcal{M}^{(13)}(\delta \otimes \delta^+)\Delta = (\mathrm{Id} \otimes \Delta)\delta;$$

the proofs of other properties are left to the readers. See also Proposition 30. The second identity of (9.5) yields

$$\delta^+ \circ p_+ = (\mathrm{Id} \otimes p_+)\delta^+$$

on T^+. Thus, we have

$$\begin{aligned}
\mathcal{M}^{(13)}(\delta \otimes \delta^+)\Delta &= \mathcal{M}^{(13)}\big(\delta \otimes (\delta^+ \circ p_+)\big)\mathsf{D} \\
&= \mathcal{M}^{(13)}\big(p_- \otimes \mathrm{Id} \otimes p_- \otimes p_+\big)\big(\mathsf{D}^- \otimes \overline{\mathsf{D}}^-\big)\mathsf{D} \\
&= \big(p_- \otimes \mathrm{Id} \otimes p_+\big)\mathcal{M}^{(13)}\big(\mathsf{D}^- \otimes \overline{\mathsf{D}}^-\big)\mathsf{D} \\
&= \big(p_- \otimes \mathrm{Id} \otimes p_+\big)(\mathrm{Id} \otimes \mathsf{D})\mathsf{D}^- \\
&= \big(\mathrm{Id} \otimes \Delta\big)\delta.
\end{aligned}$$

(d) Recall the explicit formula for the map $\upharpoonleft\!\downharpoonright_{\mathsf{e}}$, from Lemma 33. It is obvious that U is stable under $\upharpoonleft\!\downharpoonright_{\mathsf{e}}$. Define

$$\upharpoonleft\!\downharpoonright(\tau_{\mathfrak{e}}^{\mathfrak{n},\mathfrak{o}}) := \sum_{\sigma \in A(\tau)} \sum_{\mathfrak{n}_\sigma, \mathfrak{e}'_{\partial\sigma}} \frac{1}{\mathfrak{e}'_{\partial\sigma}!} \binom{\mathfrak{n}}{\mathfrak{n}_\sigma} (\tau/\sigma)^{\mathfrak{n}-\mathfrak{n}_\sigma, \mathfrak{o}|_{\tau\setminus\sigma}}_{\mathfrak{e}+\mathfrak{e}'_{\partial\sigma}} \otimes \sigma_{\mathfrak{e}}^{\mathfrak{n}_\sigma+\mathfrak{e}'_{\partial\sigma}, \mathfrak{o}},$$

where $A(\tau) := \{P_e\tau\}_{e\in E_\tau}$ – where $P_e\tau$ is a connected component of the graph $\tau/\{e\}$ containing the root of τ; see (6.5). Comparing this with the definition of Δ^+, it is not difficult to show the equality

$$\mathcal{M}^{(13)}(\delta \otimes \delta) \upharpoonleft\!\downharpoonright = (\mathrm{Id} \otimes \upharpoonleft\!\downharpoonright)\delta$$

proceeding as in the proof of point (c). Note that the contracted tree τ/σ is always planted. Let p_{e} be the canonical projection on the set of planted trees η with

$$\mathfrak{n}(\rho_\eta) = 0, \quad e(e_\eta) = \mathsf{e},$$

where e_η is the only one edge leaving the root ρ_η, and let c be the map sending the tree of the form $\mathcal{I}_n(\tau)$ to τ. Then,

$$\upharpoonleft\!\downharpoonright_{\mathsf{e}}|_U = (c \circ p_{\mathsf{e}} \otimes \mathrm{Id}) \upharpoonleft\!\downharpoonright$$

on U. Since it is elementary to show

$$(\mathrm{Id} \otimes c)\delta = \delta \circ c,$$
$$(\mathrm{Id} \otimes p_{\mathrm{e}})\delta = \delta \circ p_{\mathrm{e}},$$

the compatibility condition follows by writing

$$\begin{aligned} \mathcal{M}^{(13)}(\delta \otimes \delta)\upharpoonleft\!\downharpoonright_{\mathrm{e}}|_U &= \mathcal{M}^{(13)}\big((\mathrm{Id} \otimes c \circ p_{\mathrm{e}})\delta \otimes \delta\big) \upharpoonleft\!\downharpoonright \\ &= \big(\mathrm{Id} \otimes c \circ p_{\mathrm{e}} \otimes \mathrm{Id}\big)\mathcal{M}^{(13)}(\delta \otimes \delta) \upharpoonleft\!\downharpoonright \\ &= \big(\mathrm{Id} \otimes c \circ p_{\mathrm{e}} \otimes \mathrm{Id}\big)\big(\mathrm{Id} \otimes \upharpoonleft\!\downharpoonright\big)\delta = (\mathrm{Id} \otimes (\upharpoonleft\!\downharpoonright_{\mathrm{e}}|_U))\delta. \qquad \blacksquare \end{aligned}$$

9.4. Some examples

– For simplicity, we consider the equation

$$(\partial_t - \Delta_x + 1)u = f(u)\zeta + g(u)(\partial_x u)^2$$

for $d = 1$ with the noise $\zeta \in \mathcal{C}^{-1-\kappa}$ for sufficiently small $\kappa > 0$. Theorem 43 yields that, for any $k \in G^-_{\mathrm{ad}}$, one has the renormalized equation

$$\begin{aligned} (\partial_t - \Delta_x + 1)u^{(k)} = \ & f(u^{(k)})\zeta + g(u^{(k)})(\partial_x u^{(k)})^2 \\ & + k(\circ) f(u^{(k)}) + k(\circ_1) f'(u^{(k)})\partial_x u^{(k)} \\ & + 2k\Big(\overset{\circ}{\underset{\bullet}{\|}}\Big) f(u^{(k)}) g(u^{(k)})\partial_x u^{(k)} \\ & + k\Big(\overset{\circ}{\underset{\circ}{|}}\Big) f(u^{(k)}) f'(u^{(k)}) + k\Big(\overset{\circ\ \circ}{\underset{\bullet}{\vee}}\Big) f^2(u^{(k)}) g(u^{(k)}). \end{aligned}$$

The double line $\|$ represents the edge with $\mathfrak{e}$-decoration $(0, 1) \in \mathbb{N} \times \mathbb{N}$. The dot $\circ_1$ represents the node with $\mathfrak{n}$-decoration $(0, 1) \in \mathbb{N} \times \mathbb{N}$. More terms are needed when $\zeta \in \mathcal{C}^{-3/2-\kappa}$.

– Table 2 shows a list of strongly conforming trees associated with the generalized (KPZ) equation (6.1), without red nodes. Fix $d = 1$ for simplicity. Fix also

$$\beta_0 = -3/2 - \kappa$$

for sufficiently small $\kappa > 0$. The dot $\bullet_1$ represents the node with $\mathfrak{n}$-decoration $(0, 1) \in \mathbb{N} \times \mathbb{N}$.

– Here are some examples of the actions of splitting map D^-. The dot ${\color{red}\bullet}(\alpha)$ represents the node with $\mathfrak{o}$-decoration $\alpha \in \mathbb{Z}[\beta_0]$:

$$\mathsf{D}^- \circ = \mathbf{1}_- \otimes \circ + \circ \otimes {\color{red}\bullet}(\beta_0),$$

$$\mathsf{D}^- \overset{\circ}{\underset{\bullet}{|}} = \mathbf{1}_- \otimes \overset{\circ}{\underset{\bullet}{|}} + \circ \otimes \overset{{\color{red}\bullet}(\beta_0)}{\underset{\bullet}{|}} + \bullet \otimes \overset{\circ}{\underset{{\color{red}\bullet}(0)}{|}} + \circ\bullet \otimes \overset{{\color{red}\bullet}(\beta_0)}{\underset{{\color{red}\bullet}(0)}{|}} + \overset{\circ}{\underset{\bullet}{|}} \otimes {\color{red}\bullet}(\beta_0 + 1).$$

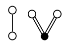
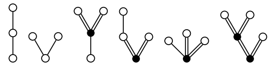
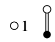
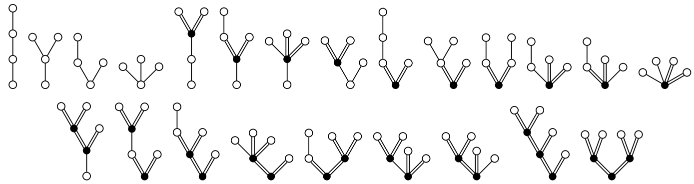
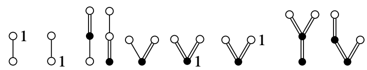
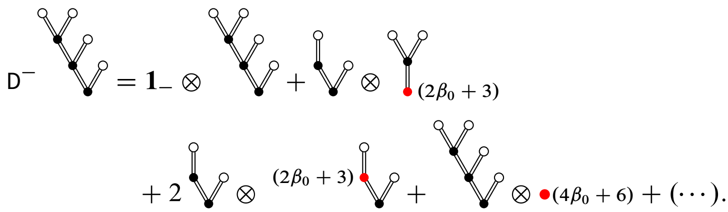
Homogeneity	Rooted decorated trees
$\beta_0 = -3/2 - \kappa$	
$2\beta_0 + 2 = -1 - 2\kappa$	
$3\beta_0 + 4 = -1/2 - 3\kappa$	
$\beta_0 + 1 = -1/2 - \kappa$	
$4\beta_0 + 6 = -4\kappa$	
$2\beta_0 + 3 = -2\kappa$	

Table 2. List of symbols of negative homogeneity for the (gKPZ) equation in the mildly singular setting where $\zeta \in \mathcal{C}^{-3/2-\kappa}$, with $\kappa > 0$ small.

For larger trees, it is inconvenient to write down all possible terms. Note that some of them vanish by the application of p_- or $k \in G_{ad}^-$. Omitting them by $(\cdots)$, one has for example



A. Summary of notations

Table 3 is a summary of the notations that we used in several sections.

B. Basics from algebra

We recall some basics of bialgebras, Hopf algebras, and comodules without proofs. See [40, 73, 78] for details. Note that, for any two algebras A and B with units $\mathbf{1}_A$ and $\mathbf{1}_B$,

Notations	Section	Meaning
$\mathcal{T} = ((T^+, \Delta^+), (T, \Delta))$	2.2	(Concrete) regularity structure.
$\mathcal{B}^+, \mathcal{B}$	2.2	Bases of T^+ and T.
$\mathcal{S}_+$	2.2	Antipode of T^+.
$G^+, \hat{g}$	2.2	Character group of T^+ and an action of $g \in G^+$ on T.
$d(\cdot,\cdot), \lvert\cdot\rvert_{\mathfrak{s}}$	2.3	Scaled metric and scaled degree.
$(T_X^+, T_X), (\mathcal{B}_X^+, \mathcal{B}_X)$	2.3	Polynomial regularity structure and their bases.
$\mathcal{D}_{(0,t)}^{\gamma,\eta}$	4.3	Singular modelled distributions on the time interval $(0,t)$.
$\mathcal{U} = ((U, \delta), (U^-, \delta^-))$	5.1	Renormalization structure.
$G^-, \tilde{k}$	5.1	Character group of U^- and an action of $k \in G^-$ on U.
${}^k\mathsf{M}$	5.2	Renormalized model.
$\tau, N_\tau, E_\tau, \rho_\tau$	6.1	Rooted tree, node set, edge set, and root.
V, V^*	6.1	Vector space spanned by rooted decorated trees and its copy space.
$\curvearrowright^{\mathsf{e}}, \curvearrowright^{\mathsf{e}}_{\mathfrak{b}}$	6.1	Grafting operator $V^* \otimes V^* \to V^*$ and its projection on U^*.
$\mathsf{M}^\zeta = (\mathsf{g}^\zeta, \Pi^\zeta)$	6.1	Canonical model associated with a smooth noise ζ.
$\upharpoonleft\!\!\vdash_{\mathsf{e}}, \upharpoonleft\!\!\vdash_{\mathsf{e}}\lvert_U$	6.3	Dual map of $\curvearrowright_{\mathsf{e}}$ and its restriction to U.
$\mathcal{S}'_- : U^- \to \mathbb{R}[U]$	7	Twisted negative antipode.
$\mathcal{SC}$	9.1	Set of all strongly conforming decorated trees.
$\mathcal{C}$	9.1	Set of all conforming decorated trees.

Table 3. List of symbols used in the main body of the text.

respectively, the tensor space $A \otimes B$ is also an algebra with the product

$$(a_1 \otimes b_1) \cdot (a_2 \otimes b_2) := (a_1 a_2) \otimes (b_1 b_2), \quad (a_1, a_2 \in A, b_1, b_2 \in B)$$

and with unit $\mathbf{1}_A \otimes \mathbf{1}_B$.

Definition. A *bialgebra* $(B, \mathcal{M}, \mathbf{1}, \triangle, \theta)$ is a 5-tuple of the following components.

- An algebra B with product $\mathcal{M} : B \otimes B \to B$, and unit $\mathbf{1}$.
- An algebra morphism $\triangle : B \to B \otimes B$ satisfying the *co-associativity*

$$(\triangle \otimes \mathrm{Id})\triangle = (\mathrm{Id} \otimes \triangle)\triangle.$$

- An algebra morphism $\theta : B \to \mathbb{R}$, satisfying

$$(\theta \otimes \mathrm{Id})\triangle = (\mathrm{Id} \otimes \theta)\triangle = \mathrm{Id},$$

where we identify $a \otimes \tau = \tau \otimes a = a\tau$ for any $a \in \mathbb{R}$ and $\tau \in B$.

The map $\triangle$ is called a *coproduct*, and the map θ is called a *counit*. An algebra morphism $S : B \to B$, such that

$$\mathcal{M}(\mathrm{Id} \otimes S)\triangle = \mathcal{M}(S \otimes \mathrm{Id})\triangle = \theta(\cdot)\mathbf{1} \tag{B.1}$$

is called an *antipode*. A bialgebra equipped with an antipode S is called a *Hopf algebra*.

The counit $\theta(\cdot)$ is traditionally denoted by $\varepsilon(\cdot)$. We use a different letter as ε already stands for a regularization parameter in this work. The following result gives a sufficient condition for a bialgebra to be a Hopf algebra. A bialgebra B is called *graded* if it is a direct sum $\bigoplus_{\lambda\in\Lambda} B_\lambda$ of vector spaces such that

- Λ is a locally finite subset of $[0, \infty)$ such that $0 \in \Lambda$ and $\Lambda + \Lambda \subset \Lambda$,
- $\mathbf{1} \in B_0$ and $B_\lambda \cdot B_\mu \subset B_{\lambda+\mu}$, for any $\lambda, \mu \in \Lambda$,
- $\triangle B_\lambda \subset \bigoplus_{\mu,\nu\in\Lambda,\,\mu+\nu=\lambda} B_\mu \otimes B_\nu$.

We call Λ a *grading* in this paper. A graded bialgebra with $B_0 = \langle \mathbf{1} \rangle$ is said to be *connected*.

Proposition 48 ([78, Exercises pages 228 and 238], [73, Proposition II.1.1 and Corollary II.3.2]). *Any connected graded bialgebra is a Hopf algebra. Moreover, one has the following properties.*

- $\theta(\mathbf{1}) = 1$ *and* $\theta(\tau) = 0$ *for any* $\tau \in \bigoplus_{\lambda>0} B_\lambda$.
- $\triangle\mathbf{1} = \mathbf{1} \otimes \mathbf{1}$, *and for any* $\tau \in B_\lambda$ *with* $\lambda > 0$,

$$\triangle\tau \in \Big\{\tau \otimes \mathbf{1} + \mathbf{1} \otimes \tau + \sum_{\substack{\mu,\nu\in\Lambda \\ \mu+\nu=\lambda,\, 0<\mu<\lambda}} B_\mu \otimes B_\nu\Big\}.$$

Based on the first assertion, we denote by $\mathbf{1}'$ the counit θ of a connected graded bialgebra. The preceding formula for $\triangle\tau$ gives an inductive formula for the antipode. For $\tau \neq \mathbf{1}$ and $\triangle\tau = \tau \otimes \mathbf{1} + \mathbf{1} \otimes \tau + \sum \tau_1 \otimes \tau_2$, one has

$$S(\tau) = -\tau - \sum S(\tau_1)\tau_2.$$

On the dual space B' of the bialgebra B, the *convolution product* is defined by

$$(f * g)\tau := (f \otimes g)\triangle\tau$$

for all $f, g \in B'$, $\tau \in B$, where we identify $a \otimes b = ab$ for any $a, b \in \mathbb{R}$. The coassociativity of $\triangle$ implies the associativity of the convolution

$$(f * g) * h = f * (g * h)$$

for all $f, g \in B'$, and the counit θ is indeed a unit of the convolution product

$$f * \theta = \theta * f = f$$

for all $f \in B'$. Hence, the triplet $(B', *, \theta)$ is a unital ring. Moreover, the subset $G \subset B'$ of algebra morphisms $g : B \to \mathbb{R}$ is stable under the convolution product. The existence of an antipode S implies that G is a group. Indeed, the inverse of $g \in G$ is given by $g^{-1} = g \circ S$. Each element of G is called a *character*, and when B is a Hopf algebra, the set G is called the *character group*.

We recall comodules and comodule bialgebras. Given an algebra A and two spaces E, F, we define on the algebraic tensor product $A \otimes E \otimes A \otimes F$ the $A \otimes E \otimes F$-valued map

$$\mathcal{M}^{(13)}\big(a_1 \otimes e \otimes a_2 \otimes f\big) := (a_1 a_2) \otimes e \otimes f.$$

Definition. Let $(B, \mathcal{M}, \mathbf{1}, \triangle, \theta)$ be a bialgebra.

- A linear space M equipped with a linear map $\delta : M \to B \otimes M$, with the properties

  $$(\mathrm{Id}_B \otimes \delta)\delta = (\triangle \otimes \mathrm{Id}_M)\delta \quad \text{and} \quad (\theta \otimes \mathrm{Id}_M)\delta = \mathrm{Id}_M,$$

  is called a left B-*comodule*. Similarly, a linear space N is called a right B-comodule if a linear map $\rho : N \to N \otimes B$ exists and satisfies

  $$(\rho \otimes \mathrm{Id}_B)\rho = (\mathrm{Id}_N \otimes \triangle)\rho \quad \text{and} \quad (\mathrm{Id}_N \otimes \theta)\rho = \mathrm{Id}_N.$$

- A bialgebra M is called a left B-*comodule bialgebra* if M is a left B-comodule by an algebra morphism $\delta : M \to B \otimes M$ such that

  $$\mathcal{M}^{(13)}(\delta \otimes \delta)\triangle_M = (\mathrm{Id} \otimes \triangle_M)\delta \quad \text{and} \quad (\mathrm{Id} \otimes \theta_M)\delta = \theta_M(\cdot)\mathbf{1},$$

  where $\triangle_M$ is a coproduct of M and θ_M is a counit of M.

Proposition 49 ([40, Proposition 2]). *Let M be a B-comodule bialgebra. If M has an antipode S_M, then*

$$\delta \circ S_M = (\mathrm{Id}_B \otimes S_M)\delta.$$

C. Technical proofs

This section is dedicated to proving Theorem 20 and Lemma 46.

C.1. Proof of Theorem 20

For any $x \in \mathbb{R} \times \mathbb{R}^d$ and $\lambda \in (0, 1]$, denote by $\varphi \mapsto \varphi_x^\lambda$ the transformation of functions on $\mathbb{R} \times \mathbb{R}^d$ defined by

$$\varphi_x^\lambda(y) := \lambda^{-d-2}\varphi\big(\lambda^{-2}(y_0 - x_0), \lambda^{-1}(y' - x')\big).$$

The following bound appears in Hairer's original paper [53]. Recall $\beta_0 = \min A$ and that p_t denotes the heat kernel of the operator $\widetilde{\mathscr{G}} = \partial_{x_0}^2 - (\Delta_{x'} - 1)^2$ on $\mathbb{R} \times \mathbb{R}^d$.

Lemma 50. *Let* $\mathsf{M} = (\Pi, g)$ *be a model over the regularity structure* $\mathcal{T}$ *and* $\boldsymbol{f} \in \mathcal{D}^\gamma(T, \mathrm{g})$ *with* $\gamma \in \mathbb{R}$*. Assume* $\beta_0 > -4$*. Then, for any Schwartz function* $\varphi \in \mathcal{S}(\mathbb{R} \times \mathbb{R}^d)$*,* $x \in \mathbb{R} \times \mathbb{R}^d$*, and* $\lambda \in (0, 1]$*, one has the bound*

$$\left|\left\langle \mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x), \varphi_x^\lambda \right\rangle\right| \le C_\varphi \|\Pi^{\mathrm{g}}\| \|\boldsymbol{f}\|_{\mathcal{D}^\gamma} \lambda^\gamma,$$

where the constant C_φ *depends on the size* $\sup_{|k|_{\mathfrak{s}}, |\ell|_{\mathfrak{s}} \le N} \|x^k \partial^\ell \varphi\|_{L^\infty(\mathbb{R} \times \mathbb{R}^d)}$ *for* $N > 0$ *large enough.*

Proof. Write $\Lambda_x := \mathbf{R}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathrm{g}} \boldsymbol{f}(x)$ to shorten notations. Using

$$p_0 = \int_0^{\lambda^4} \widetilde{\mathscr{G}} p_t dt + p_{\lambda^4}$$

and the symmetry of $\widetilde{\mathscr{G}}$,

$$\begin{aligned}
\langle \Lambda_x, \varphi_x^\lambda \rangle &= \int_{\mathbb{R} \times \mathbb{R}^d} \int_0^{\lambda^4} \left\langle \Lambda_x, \widetilde{\mathscr{G}}_y p_t(y, \cdot) \right\rangle \varphi_x^\lambda(y) dt\, dy \\
&\quad + \int_{\mathbb{R} \times \mathbb{R}^d} \left\langle \Lambda_x, p_{\lambda^4}(y, \cdot) \right\rangle \varphi_x^\lambda(y) dy \\
&= \int_{\mathbb{R} \times \mathbb{R}^d} \int_0^{\lambda^4} \left\langle \Lambda_x, p_t(y, \cdot) \right\rangle \widetilde{\mathscr{G}}_y \varphi_x^\lambda(y) dt\, dy \\
&\quad + \int_{\mathbb{R} \times \mathbb{R}^d} \left\langle \Lambda_x, p_{\lambda^4}(y, \cdot) \right\rangle \varphi_x^\lambda(y) dy =: (A) + (B).
\end{aligned}$$

Using the properties of models as in Proposition 2, one has

$$\left|\langle \Lambda_x, p_t(y, \cdot) \rangle\right| \lesssim t^{\frac{\gamma}{4}} + \sum_{\beta \in [\beta_0, \gamma)} t^{\frac{\beta}{4}} d(y, x)^{\gamma - \beta}.$$

This implies $|(B)| \lesssim \lambda^\gamma$. Moreover, since $\beta_0 > -4$, one has

$$\begin{aligned}
|(A)| &\lesssim \sum_{\beta \in [\beta_0, \gamma]} \int_0^{\lambda^4} t^{\frac{\beta}{4}} dt \int_{\mathbb{R} \times \mathbb{R}^d} d(y, x)^{\gamma - \beta} \, |\widetilde{\mathscr{G}}_y \varphi_x^\lambda(y)| \, dy \\
&\lesssim \sum_{\beta \in [\beta_0, \gamma]} \lambda^{\beta + 4} \lambda^{\gamma - \beta - 4} \lesssim \lambda^\gamma.
\end{aligned}$$

■

We prove a fundamental fact on the connection of modelled distributions. For any interval $I \subset \mathbb{R}$, denote by $\mathcal{D}_I^\gamma(T, \mathrm{g})$ the set of functions $\boldsymbol{f} : I \times \mathbb{R}^d \to T_{<\gamma}$ which satisfies the bounds of $[\![\boldsymbol{f}]\!]_{\mathcal{D}_I^\gamma}$ and $\|\boldsymbol{f}\|_{\mathcal{D}_I^\gamma}$ as in Definition 3 with $\mathbb{R} \times \mathbb{R}^d$ replaced by $I \times \mathbb{R}^d$.

Lemma 51. *Let* $\mathsf{M} = (g, \Pi)$ *be a model over* $\mathcal{T}$*, and let* $a < b < c$*. If* $\boldsymbol{f} : [a, c] \times \mathbb{R}^d \to T_{<\gamma}$ *satisfies the bounds of* $|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{[a,b]}^\gamma}$ *and* $|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{[b,c]}^\gamma}$*, then* $\boldsymbol{f} \in \mathcal{D}_{[a,c]}^\gamma(T, \mathrm{g})$ *and one has*

$$|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{[a,c]}^\gamma} \lesssim (1 + \|\hat{\mathrm{g}}\|_\gamma) \left(|\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{[a,b]}^\gamma} + |\!|\!| \boldsymbol{f} |\!|\!|_{\mathcal{D}_{[b,c]}^\gamma} \right).$$

Proof. It is sufficient to show the bound of $\|\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\|_\beta$ for $x \in [a,b]\times\mathbb{R}^d$ and $y \in [b,c]\times\mathbb{R}^d$. Setting $z = (b, y')$, we have

$$\begin{aligned}
\big\|\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yx}}\boldsymbol{f}(x)\big\|_\beta &\le \big\|\boldsymbol{f}(y) - \widehat{\mathrm{g}_{yz}}\boldsymbol{f}(z)\big\|_\beta + \big\|\widehat{\mathrm{g}_{yz}}\big(\boldsymbol{f}(z) - \widehat{\mathrm{g}_{zx}}\boldsymbol{f}(x)\big)\big\|_\beta \\
&\le \|\boldsymbol{f}\|_{\mathcal{D}^\gamma_{[b,c]}} d(y,z)^{\gamma-\beta} \\
&\quad + \|\hat{\mathrm{g}}\|_\gamma \|\boldsymbol{f}\|_{\mathcal{D}^\gamma_{[a,b]}} \sum_{\alpha\in[\beta,\gamma)} d(y,z)^{\alpha-\beta} d(z,x)^{\gamma-\alpha} \\
&\lesssim \big(1 + \|\hat{\mathrm{g}}\|_\gamma\big)\big(|||\boldsymbol{f}|||_{\mathcal{D}^\gamma_{[a,b]}} + |||\boldsymbol{f}|||_{\mathcal{D}^\gamma_{[b,c]}}\big) d(y,x)^{\gamma-\beta}.
\end{aligned}$$

■

We recall now from J. Martin's work [74, Theorem 5.3.16] the existence of a "Whitney extension" map on locally defined modelled distributions.

Theorem 52. *Let* $\mathsf{M} = (g, \Pi)$ *be a model over* $\mathcal{T}$ *with a regular product* $\star$ *satisfying Assumption A. Then, there exists a continuous linear operator*

$$E : \mathcal{D}^\gamma_{(-\infty,t]}(T, \mathrm{g}) \to \mathcal{D}^\gamma(T, \mathrm{g})$$

such that $(E\boldsymbol{f})|_{(-\infty,t]\times\mathbb{R}^d} = \boldsymbol{f}$*, and the bound*

$$|||E\boldsymbol{f}|||_{\mathcal{D}^\gamma} \le C\,|||\boldsymbol{f}|||_{\mathcal{D}^\gamma_{(-\infty,t]}}$$

holds for a positive constant $C = C(\|\hat{\mathrm{g}}\|_\gamma)$ *depending polynomially on* $\|\hat{\mathrm{g}}\|_\gamma$ *and independent of* $t > 0$ *and* $\boldsymbol{f} \in \mathcal{D}^\gamma_{(-\infty,t]}(T, \mathrm{g})$*. A similar result holds for the modelled distributions defined on* $[t,\infty)\times\mathbb{R}^d$.

Proof. For simplicity, we consider $t = 0$. It is sufficient to construct the continuous linear extension operator $\widetilde{E} : \mathcal{D}^\gamma_{(-\infty,0]}(T, \mathrm{g}) \to \mathcal{D}^\gamma_{(-\infty,1]}(T, \mathrm{g})$. Indeed, once we pick $\chi \in C^\infty(\mathbb{R}\times\mathbb{R}^d)$ which depends only on x_0 such that $\chi(x_0, x') = 1$ when $x_0 \le 0$ and $\chi(x_0, x') = 0$ when $x_0 \ge 1$, and set

$$\boldsymbol{\chi}(x) = \sum_{|k|_{\mathfrak{s}} < \gamma - \beta_0} \frac{\partial^k \chi(x)}{k!} X^k,$$

then by Proposition 8, we can define the modelled distribution on $\mathbb{R}\times\mathbb{R}^d$ by

$$E\boldsymbol{f} := \mathcal{Q}_{<\gamma}\big(\boldsymbol{\chi} \star (\widetilde{E}\boldsymbol{f})\big).$$

Then, the operator $E : \mathcal{D}^\gamma_{(-\infty,0]}(T, \mathrm{g}) \to \mathcal{D}^\gamma(T, \mathrm{g})$ satisfies the desired properties.

We construct the extension operator $\widetilde{E} : \mathcal{D}^\gamma_{(-\infty,0]}(T, \mathrm{g}) \to \mathcal{D}^\gamma_{(-\infty,1]}(T, \mathrm{g})$. Let

$$h(x_0, x') = h_{x_0}(x')$$

be the kernel of the operator $e^{x_0\Delta_{x'}}$ with $x_0 > 0$, and define the function

$$\boldsymbol{h}(x) = \sum_{|k|_{\mathfrak{s}} < \gamma - \beta_0} \frac{\partial^k h(x)}{k!} X^k$$

on $(0,\infty)\times\mathbb{R}^d$. Then, we have the properties

$$\int_{\mathbb{R}^d} \boldsymbol{h}(x)dx' = \mathbf{1}, \qquad \int_{\mathbb{R}^d} |\partial^k h(x_0,x')||x'|^\alpha \lesssim x_0^{\frac{\alpha-|k|_{\mathfrak{s}}}{2}} \tag{C.1}$$

for any $\alpha\ge 0$. For any $\boldsymbol{f}\in\mathcal{D}^\gamma_{(-\infty,0]}(T,\mathrm{g})$, we set $(\widetilde{E}\boldsymbol{f})(x) := \boldsymbol{f}(x)$ if $(-\infty,0]\times\mathbb{R}^d$ and

$$(\widetilde{E}\boldsymbol{f})(x) := \mathcal{Q}_{<\gamma}\int_{\mathbb{R}^d} \boldsymbol{h}\big(x-(0,y')\big)\star\widehat{\mathrm{g}_{x(0,y')}}f(0,y')\,dy'$$

if $x\in(0,\infty)\times\mathbb{R}^d$. We first prove

$$\|(\widetilde{E}\boldsymbol{f})(x)-\widehat{\mathrm{g}_{x(0,x')}}\boldsymbol{f}(0,x')\|_\beta \lesssim \|\widehat{\mathrm{g}}\|\|\boldsymbol{f}\|_{\mathcal{D}^\gamma}\, x_0^{\frac{\gamma-\beta}{2}} \tag{C.2}$$

for any $\beta<\gamma$ and $x=(x_0,x')\in[0,1]\times\mathbb{R}^d$. Since

$$\begin{aligned}&(\widetilde{E}\boldsymbol{f})(x)-\widehat{\mathrm{g}_{x(0,x')}}\boldsymbol{f}(0,x')\\ &\quad=\int_{\mathbb{R}^d}\boldsymbol{h}(x_1,x'-y')\star\widehat{\mathrm{g}_{x(0,x')}}\big\{\widehat{\mathrm{g}_{(0,x')(0,y')}}\boldsymbol{f}(0,y')-\boldsymbol{f}(0,x')\big\}dy'\end{aligned}$$

by the first property of (C.1), we have (C.2) as follows:

$$\begin{aligned}&\|(\widetilde{E}\boldsymbol{f})(x)-\widehat{\mathrm{g}_{x(0,x')}}\boldsymbol{f}(0,x')\|_\beta\\
&\le\sum_{|k|_{\mathfrak{s}}<\gamma-\beta_0}\int_{\mathbb{R}^d}|\partial^k h(x_0,x'-y')|\Big\|\widehat{\mathrm{g}_{x(0,x')}}\big\{\widehat{\mathrm{g}_{(0,x')(0,y')}}\boldsymbol{f}(0,y')-\boldsymbol{f}(0,x')\big\}\Big\|_{\beta-|k|_{\mathfrak{s}}}dy'\\
&\le\|\widehat{\mathrm{g}}\|\|\boldsymbol{f}\|_{\mathcal{D}^\gamma}\sum_{|k|_{\mathfrak{s}}<\gamma-\beta_0}\ \sum_{\alpha\in[\beta-|k|_{\mathfrak{s}},\gamma)}\int_{\mathbb{R}^d}|\partial^k h(x_0,x'-y')|\,x_0^{\frac{\alpha-\beta+|k|_{\mathfrak{s}}}{2}}|x'-y'|^{\gamma-\alpha}dy'\\
&\lesssim\|\widehat{\mathrm{g}}\|\|\boldsymbol{f}\|_{\mathcal{D}^\gamma}\sum_{|k|_{\mathfrak{s}}<\gamma-\beta_0}\ \sum_{\alpha\in[\beta-|k|_{\mathfrak{s}},\gamma)} x_0^{\frac{\alpha-\beta+|k|_{\mathfrak{s}}}{2}}\,x_0^{\frac{\gamma-\alpha-|k|_{\mathfrak{s}}}{2}}\lesssim\|\widehat{\mathrm{g}}\|\|\boldsymbol{f}\|_{\mathcal{D}^\gamma}\,x_0^{\frac{\gamma-\beta}{2}}.\end{aligned}$$

The bound of $\|(\widetilde{E}\boldsymbol{f})(x)\|_\beta$ on $x\in[0,1]\times\mathbb{R}^d$ follows from (C.2) and the bound of $\boldsymbol{f}(0,x')$. For the bound of $\|(\widetilde{E}\boldsymbol{f})(y)-\widehat{\mathrm{g}_{yx}}(\widetilde{E}\boldsymbol{f})(x)\|_\beta$, it is sufficient to consider the case that $x,y\in[0,1]\times\mathbb{R}^d$ by Lemma 51. If $x_0^{\frac12}\wedge y_0^{\frac12}\le d(x,y)$, we can bound above $\|(\widetilde{E}\boldsymbol{f})(y)-\widehat{\mathrm{g}_{yx}}(\widetilde{E}\boldsymbol{f})(x)\|_\beta$ by

$$\begin{aligned}&\Big\|(\widetilde{E}\boldsymbol{f})(y)-\widehat{\mathrm{g}_{y(0,y')}}\boldsymbol{f}(0,y')\Big\|_\beta+\Big\|\widehat{\mathrm{g}_{y(0,y')}}\big\{\boldsymbol{f}(0,y')-\widehat{\mathrm{g}_{(0,y')(0,x')}}\boldsymbol{f}(0,x')\big\}\Big\|_\beta\\
&\quad+\Big\|\widehat{\mathrm{g}_{yx}}\big\{\widehat{\mathrm{g}_{x(0,x')}}\boldsymbol{f}(0,x')-(\widetilde{E}\boldsymbol{f})(x)\big\}\Big\|_\beta,\end{aligned}$$

and using (C.2), we obtain the desired bound. If $d(x,y)<x_0^{\frac12}\wedge y_0^{\frac12}$, by using the formula

$$\begin{aligned}\boldsymbol{h}(y)-\widehat{\mathrm{g}_{yx}}\boldsymbol{h}(x)&=\sum_{|k|_{\mathfrak{s}}<\gamma-\beta_0}\frac{X^k}{k!}\Big(\partial^k h(y)-\sum_{|\ell|_{\mathfrak{s}}<\gamma-\beta_0-|k|_{\mathfrak{s}}}\frac{(y-x)^\ell}{\ell!}\partial^{k+\ell}h(x)\Big)\\
&=\sum_{|k|_{\mathfrak{s}}<\gamma-\beta_0}\frac{X^k}{k!}\sum_{|n|_{\mathfrak{s}}\ge\gamma-\beta_0-|k|_{\mathfrak{s}}}\frac{(y-x)^n}{n!}\int_{[0,1]^{d+1}}\partial^{k+n}h(x_t)m_n(dt)\end{aligned}$$

obtained from Lemma 18, where n runs over a finite set, we decompose

$$\begin{aligned}
&(\widetilde{E}\boldsymbol{f})(y) - \widehat{\mathrm{g}_{yx}}(\widetilde{E}\boldsymbol{f})(x) \\
&= \mathcal{Q}_{<\gamma} \int_{\mathbb{R}^d} \{\boldsymbol{h}(y-(0,z')) - \widehat{\mathrm{g}_{yx}}\boldsymbol{h}(x-(0,z'))\} \star \widehat{\mathrm{g}_{y(0,z')}}\boldsymbol{f}(0,z')dz' \\
&= \mathcal{Q}_{<\gamma} \sum_{\substack{|k|_{\mathfrak{s}}<\gamma-\beta_0 \\ |n|_{\mathfrak{s}}\ge\gamma-\beta_0-|k|_{\mathfrak{s}}}} \frac{(y-x)^n}{k!n!} \int_{\mathbb{R}^d}\int_{[0,1]^{d+1}} \partial^{k+n}h(x_t-(0,z'))m_n(dt)X^k \\
&\quad \star \widehat{\mathrm{g}_{y(0,z')}}\boldsymbol{f}(0,z')dz'.
\end{aligned}$$

Thus, it is sufficient to show that

$$\left\| \int_{\mathbb{R}^d} \partial^{k+n}h(x_t-(0,z'))\widehat{\mathrm{g}_{y(0,z')}}\boldsymbol{f}(0,z')dz' \right\|_{\beta-|k|_{\mathfrak{s}}} \lesssim d(x,y)^{\gamma-\beta-|n|_{\mathfrak{s}}}. \tag{C.3}$$

For this purpose, we decompose

$$\begin{aligned}
&\left\| \int_{\mathbb{R}^d} \partial^{k+n}h(x_t-(0,z'))\widehat{\mathrm{g}_{y(0,z')}}\boldsymbol{f}(0,z')dz' \right\|_{\beta-|k|_{\mathfrak{s}}} \\
&\le \left\| \int_{\mathbb{R}^d} \partial^{k+n}h(x_t-(0,z'))\widehat{\mathrm{g}_{y(0,x_t')}}\{\widehat{\mathrm{g}_{(0,x_t')(0,z')}}\boldsymbol{f}(0,z') - \boldsymbol{f}(0,x_t')\}dz' \right\|_{\beta-|k|_{\mathfrak{s}}} \\
&\le \|\hat{\mathrm{g}}\|_\gamma \|\boldsymbol{f}\|_{\mathcal{D}_0^\gamma} \sum_{\alpha\in[\beta-|k|_{\mathfrak{s}},\gamma)} \int_{\mathbb{R}^d} \left|\partial^{k+n}h(x_t-(0,z'))\right| d(x,y)^{\alpha-\beta+|k|_{\mathfrak{s}}} |x_t'-z'|^{\gamma-\alpha} dz' \\
&\lesssim \|\hat{\mathrm{g}}\|_\gamma \|\boldsymbol{f}\|_{\mathcal{D}_0^\gamma} \sum_{\alpha\in[\beta-|k|_{\mathfrak{s}},\gamma)} (x_0\wedge y_0)^{\frac{\gamma-\alpha-|k|_{\mathfrak{s}}-|n|_{\mathfrak{s}}}{2}} d(x,y)^{\alpha-\beta+|k|_{\mathfrak{s}}}.
\end{aligned}$$

Since $\gamma-\alpha-|k|_{\mathfrak{s}}-|n|_{\mathfrak{s}} \le -\alpha+\beta_0 \le 0$ under the condition that $|n|_{\mathfrak{s}} \ge \gamma-\beta_0-|k|_{\mathfrak{s}}$ and $\alpha \in A$, we obtain the desired bound (C.3). ∎

Corollary 53. *Let* $\mathsf{M} = (g,\Pi)$ *be a model over* $\mathcal{T}$ *with a regular product* $\star$ *satisfying Assumption A. For any* $\boldsymbol{f} \in \mathcal{D}^{\gamma,\eta}(T,\mathrm{g})$ *and* $a>0$*, the restriction* $\boldsymbol{f}|_{[a,\infty)\times\mathbb{R}^d}$ *satisfies the bound*

$$\|\boldsymbol{f}|_{[a,\infty)\times\mathbb{R}^d}\|_{\mathcal{D}^\gamma_{[a,\infty)}} \le Ca^{\frac{\eta\wedge\beta_0-\gamma}{2}} \|\boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}}$$

for a positive constant $C = C(\|\hat{\mathrm{g}}\|_\gamma)$ *depending polynomially on* $\|\hat{\mathrm{g}}\|_\gamma$*, independent of* a *and* $\boldsymbol{f}$*. Therefore, by writing* $\mathbf{R}^{\mathsf{M}}_a := \mathbf{R}^{\mathsf{M}} \circ E_a$ *for the extension map* E_a *from* $\mathcal{D}^\gamma_{[a,\infty)}(T,\mathrm{g})$ *to* $\mathcal{D}^\gamma(T,\mathrm{g})$*, for any Schwartz function* $\varphi \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^d)$*,* $x\in\mathbb{R}\times\mathbb{R}^d$*, and* $\lambda\in(0,1]$*, one has the bounds*

$$\left|\langle \mathbf{R}^{\mathsf{M}}_a\boldsymbol{f} - \Pi^{\mathrm{g}}_x E_a\boldsymbol{f}(x), \varphi^\lambda_x\rangle\right| \le C'\|\boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}} a^{\frac{\eta\wedge\beta_0-\gamma}{2}}\lambda^\gamma$$

and

$$\left|\langle \mathbf{R}^{\mathsf{M}}_a\boldsymbol{f}, \varphi^\lambda_x\rangle\right| \le C'\|\boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}} a^{\frac{\eta\wedge\beta_0-\gamma}{2}}\lambda^\gamma \tag{C.4}$$

for a positive constant C' *depending polynomially on* $\|\mathrm{g}\|_\gamma + \|\Pi^{\mathrm{g}}\|_\gamma$*, independent of* a *and* $\boldsymbol{f}$*.*

Proof. It is sufficient to show the bound of $\|\boldsymbol{f}(y)-\widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\|_\beta$ for any $\beta<\gamma$ and $x,y\in[a,\infty)\times\mathbb{R}^d$ such that $d(x,y)\geq\omega(x,y)$. We decompose

$$\begin{aligned}\big\|\boldsymbol{f}(y)-\widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big\|_\beta &\leq \big\|\boldsymbol{f}(y)\big\|_\beta+\big\|\widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big\|_\beta\\ &\lesssim [\![\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}}\Big(a^{\frac{(\eta-\beta)\wedge 0}{2}}+\sum_{\beta\leq\alpha<\gamma}d(x,y)^{\alpha-\beta}a^{\frac{(\eta-\alpha)\wedge 0}{2}}\Big).\end{aligned}$$

Since $d(x,y)\geq\omega(x,y)\geq a$, we have

$$\begin{aligned}\big\|\boldsymbol{f}(y)-\widehat{\mathsf{g}_{yx}}\boldsymbol{f}(x)\big\|_\beta &\lesssim [\![\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}}d(x,y)^{\gamma-\beta}\sum_{\beta\leq\alpha<\gamma}d(x,y)^{\alpha-\gamma}a^{\frac{(\eta-\alpha)\wedge 0}{2}}\\ &\lesssim [\![\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}}d(x,y)^{\gamma-\beta}a^{\frac{\eta\wedge\beta_0-\gamma}{2}}.\end{aligned}$$

Thus, the modelled distribution $\boldsymbol{f}$ restricted to $[a,\infty)\times\mathbb{R}^d$ has the norm of size $a^{\frac{\eta\wedge\beta_0-\gamma}{2}}$. ∎

We turn to the proof of the reconstruction theorem for singular modelled distributions.

Proof of Theorem 20. The proof is just an analogue of the proof of Proposition 6.9 in Hairer's seminal work [53], so we omit a number of details. The only difference is that $p_t(x,\cdot)$ is not compactly supported. For the sake of generality, for any Schwartz function $\varphi\in\mathcal{S}(\mathbb{R}\times\mathbb{R}^d)$, $x\in\mathbb{R}\times\mathbb{R}^d$, and $\lambda\in(0,1]$, we prove the bound

$$\big|\big\langle\mathbf{R}^{\mathsf{M}}\boldsymbol{f}-\Pi_x^{\mathsf{g}}\boldsymbol{f}(x),\varphi_x^\lambda\big\rangle\big|\lesssim |\!|\!|\boldsymbol{f}|\!|\!|_{\mathcal{D}^\gamma}\omega(x)^{\eta\wedge\beta_0-\gamma}\lambda^\gamma \tag{C.5}$$

for any $\lambda\in(0,\omega(x)]$, and the bound

$$\big|\big\langle\mathbf{R}^{\mathsf{M}}\boldsymbol{f},\varphi_x^\lambda\big\rangle\big|\lesssim |\!|\!|\boldsymbol{f}|\!|\!|_{\mathcal{D}^\gamma}\lambda^\eta \tag{C.6}$$

for any $\lambda\in(0,1]$. To lighten notation, we omit the proportional constants depending on the model M. By linearity, we can assume that $\boldsymbol{f}=0$ on $(-\infty,0]\times\mathbb{R}^d$. For any $a>0$, we consider the distributions $\mathbf{R}_a^{\mathsf{M}}\boldsymbol{f}:=\mathbf{R}^{\mathsf{M}}(E_a\boldsymbol{f})$ defined in Corollary 53. Because of the local property of the reconstruction operator, Corollary 5, these distributions are compatible over all $a>0$ in the sense that $\langle\mathbf{R}_a^{\mathsf{M}}\boldsymbol{f},\varphi\rangle=\langle\mathbf{R}_b^{\mathsf{M}}\boldsymbol{f},\varphi\rangle$ for any $0<a<b$ and Schwartz functions φ supported on $[b,\infty)\times\mathbb{R}^d$, so the quantity $\langle\widetilde{\mathbf{R}}^{\mathsf{M}}\boldsymbol{f},\varphi\rangle$ is defined for any φ supported on $(0,\infty)\times\mathbb{R}^d$. Since $\boldsymbol{f}$ vanishes on $\mathbb{R}\times\mathbb{R}_0^d$, one defines $\langle\widetilde{\mathbf{R}}^{\mathsf{M}}\boldsymbol{f},\varphi\rangle=0$ if φ is supported on $(-\infty,0)\times\mathbb{R}^d$. To consider arbitrary φ, fix a family $\{\phi_{n,k}\}_{n\in\mathbb{N},k\in\mathbb{Z}^d}$ of functions of the forms

$$\phi_{n,k}=2^{-n(d+2)}\phi_{x_{n,k}}^{2^{-n}},\quad x_{n,k}=\big(2^{-2n},2^{-n}k\big)\in\mathbb{R}\times\mathbb{R}^d,$$

where ϕ is a smooth function supported on $\{x\in\mathbb{R}\times\mathbb{R}^d;\,d(0,x)<1\}$, and such that $\sum_{n,k}\phi_{n,k}(x)=1$, if $0<x_0<1/2$. To show the bound (C.6), fixing an integer n_0 such that $2^{-n_0}\simeq\lambda$, and setting $\widetilde{\phi}_{n_0}=1-\sum_{n\geq n_0,k\in\mathbb{Z}^d}\phi_{n,k}$, we decompose

$$\big\langle\widetilde{\mathbf{R}}^{\mathsf{M}}\boldsymbol{f},\varphi_x^\lambda\big\rangle=\sum_{n\geq n_0,k\in\mathbb{Z}^d}\big\langle\widetilde{\mathbf{R}}^{\mathsf{M}}\boldsymbol{f},\varphi_x^\lambda\,\phi_{n,k}\big\rangle+\big\langle\widetilde{\mathbf{R}}^{\mathsf{M}}\boldsymbol{f},\varphi_x^\lambda\,\widetilde{\phi}_{n_0}\big\rangle.$$

For the second term, since $\widetilde{\phi}_{n_0}$ is supported in $\{y \in \mathbb{R} \times \mathbb{R}^d; y_0 \geq a_{n_0} \simeq 2^{-2n_0}\}$ and $\varphi_x^\lambda \widetilde{\phi}_{n_0} = \psi_x^\lambda$ for some Schwartz function ψ which is uniform over λ, x, n_0, by (C.4), one has

$$\left|\langle \widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f}, \varphi_x^\lambda \widetilde{\phi}_{n_0}\rangle\right| \lesssim 2^{-n_0(\eta\wedge\beta_0-\gamma)}\lambda^\gamma ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}} \simeq 2^{-n_0(\eta\wedge\beta_0)} ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}.$$

For the first term, since $\varphi_x^\lambda \phi_{n,k}$ is supported in $[2^{-2n}, \infty) \times \mathbb{R}^d$ and roughly equal to

$$2^{-n(d+2)}\varphi_x^\lambda(x_{n,k})\phi_{x_{n,k}}^{2^{-n}},$$

one has

$$\begin{aligned}\left|\langle \widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f}, \varphi_x^\lambda \phi_{n,k}\rangle\right| &\lesssim 2^{-n(d+2)}|\varphi_x^\lambda(x_{n,k})|2^{-n(\eta\wedge\beta_0-\gamma)}(2^{-n})^\gamma ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}} \\ &\lesssim 2^{-n(d+2)}|\varphi_x^\lambda(x_{n,k})|2^{-n(\eta\wedge\beta_0)} ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}.\end{aligned}$$

The sum $2^{-nd}\sum_{k\in\mathbb{Z}^d}|\varphi_x^\lambda(x_{n,k})|$ is roughly bounded by $\int_{\mathbb{R}^d}|\varphi^\lambda(x)|dx' \lesssim \lambda^{-2}$. Hence,

$$\begin{aligned}\sum_{n\geq n_0, k\in\mathbb{Z}^d}\left|\langle \widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f}, \varphi_x^\lambda \phi_{n,k}\rangle\right| &\lesssim ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}\lambda^{-2}\sum_{n\geq n_0}2^{-n(\eta\wedge\beta_0+2)} \\ &\lesssim ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}\lambda^{-2}2^{-n_0(\eta\wedge\beta_0+2)} \lesssim ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}\lambda^{\eta\wedge\beta_0}.\end{aligned}$$

Here, we use the assumption $\eta \wedge \beta_0 > -2$.

The proof of (C.5) is almost the same with some modifications. We fix the same n_0 as above and decompose φ_x^λ into $\sum_{n\geq n_0, k\in\mathbb{Z}^d}\varphi_x^\lambda\phi_{n,k} + \varphi_x^\lambda\widetilde{\phi}_{n_0}$. For the dual with the second term, by taking $a \simeq x_0$ in (C.5), one has

$$\left|\langle \widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x), \varphi_x^\lambda \widetilde{\phi}_{n_0}\rangle\right| \lesssim \omega(x)^{\eta\wedge\beta_0-\gamma}\lambda^\gamma ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}.$$

For the dual with the remaining terms, one decomposes

$$\begin{aligned}\widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f} - \Pi_x^{\mathsf{g}}\boldsymbol{f}(x) &= \left(\widetilde{\mathbf{R}}^{\mathsf{M}} \boldsymbol{f} - \Pi_{x_{n,k}}^{\mathsf{g}}\boldsymbol{f}(x_{n,k})\right) + \Pi_{x_{n,k}}^{\mathsf{g}}\boldsymbol{f}(x_{n,k}) - \Pi_{x_{n,k}}^{\mathsf{g}}\left(\widehat{\mathsf{g}_{x_{n,k}x}}\boldsymbol{f}(x)\right) \\ &=: (a) + (b) + (c).\end{aligned}$$

For (a), one has

$$\left|\langle (a), \varphi_x^\lambda\phi_{n,k}\rangle\right| \lesssim 2^{-n(d+2)}|\varphi_x^\lambda(x_{n,k})|2^{-n(\eta\wedge\beta_0)}\|\Pi^{\mathsf{g}}\| ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}$$

as before. The sum $2^{-nd}\sum_{k\in\mathbb{Z}^d}|\varphi_x^\lambda(x_{n,k})|$ is roughly bounded by $\int_{\mathbb{R}^d}|\varphi^\lambda(x)|dx' \lesssim \lambda^{-2}|\lambda^{-2}x_0|^{-N}$ for any $N > 0$. By picking $N = \frac{\gamma-\eta\wedge\beta_0}{2}$, one has

$$\begin{aligned}\sum_{n\geq n_0, k\in\mathbb{Z}^d}\left|\langle (a), \varphi_x^\lambda\phi_{n,k}\rangle\right| &\lesssim \lambda^{-2}|\lambda^{-2}x_0|^{\frac{\eta\wedge\beta_0-\gamma}{2}}\sum_{n\geq n_0}2^{-n(\eta\wedge\beta_0+2)} ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}} \\ &\lesssim \lambda^{-2}|\lambda^{-2}x_0|^{\frac{\eta\wedge\beta_0-\gamma}{2}}2^{-n_0(\eta\wedge\beta_0+2)} ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}} \\ &\lesssim \omega(x)^{\eta\wedge\beta_0-\gamma}\lambda^\gamma ||| \boldsymbol{f} |||_{\mathcal{D}^{\gamma,\eta}}.\end{aligned}$$

Using the bound $\|f(x)\|_\beta \lesssim |x_0|^{(\eta-\beta)/2\wedge 0}$, one gets the same bounds as above for (b) and (c).

It remains to show the uniqueness of $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} \in \mathcal{C}^{\eta\wedge\beta_0}$ satisfying (4.6). We start from the identity

$$\big|\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\rangle\big| \lesssim \omega(x)^{\eta\wedge\beta_0-\gamma} t^{\frac{\gamma}{4}}$$

satisfied uniformly in $x \in \mathbb{R}\times\mathbb{R}^d$ such that $\omega(x)^4 \geq t$ by any other reconstruction operator $(\mathbf{R}^{\mathsf{M}})'$. Since $\mathbf{R}^{\mathsf{M}}\boldsymbol{f}, (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f} \in \mathcal{C}^{\eta\wedge\beta_0}$, we also have

$$\big|\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\rangle\big| \lesssim t^{\frac{\eta\wedge\beta_0}{4}}$$

uniformly in $x \in \mathbb{R}\times\mathbb{R}^d$ and $0 < t \leq 1$. From these bounds, for any $\varepsilon \in (0,\gamma)$, we can show that

$$\big|\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\rangle\big| \lesssim \omega(x)^{\eta\wedge\beta_0-\varepsilon} t^{\frac{\varepsilon}{4}}$$

uniformly in $x \in \mathbb{R}\times\mathbb{R}^d$ and $0 < t \leq 1$. If one chooses small $\varepsilon > 0$ such that $\eta\wedge\beta_0 - \varepsilon > -2$, since $\int_{\mathbb{R}\times\mathbb{R}^d} \omega(x)^{\eta\wedge\beta_0-\varepsilon}\varphi(x)dx < \infty$ for any Schwartz function $\varphi \in \mathcal{S}(\mathbb{R}\times\mathbb{R}^d)$, one has

$$\big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - \big(\mathbf{R}^{\mathsf{M}}\big)'\boldsymbol{f}, \varphi\big\rangle = \lim_{t\to 0}\int \big\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f} - (\mathbf{R}^{\mathsf{M}})'\boldsymbol{f}, p_t(x,\cdot)\big\rangle\varphi(x)dx = \lim_{t\to 0} O\big(t^{\frac{\varepsilon}{4}}\big) = 0. \ \blacksquare$$

In the end of this section, we provide a sketch of the proof of Theorem 17′.

Proof of Theorem 17′. The proof is carried out by a method similar to the proof of Theorem 17 with modifications on the bounds of $\mathcal{J}^{\mathsf{M}}$ and $\mathcal{N}^{\mathsf{M}}$ terms. Similarly to the proof of Theorem 17, we use the decomposition $\mathcal{J}^{\mathsf{M}} = \int_0^1 \mathcal{J}_r^{\mathsf{M}} dr$ and $\mathcal{N}^{\mathsf{M}} = \int_0^1 \mathcal{N}_r^{\mathsf{M}} dr$ and write $(\tau)_{X^n}$ for the X^n-component of $\tau \in T$. As for the bound of $[\![\mathcal{K}^{\mathsf{M}}\boldsymbol{f}]\!]_{\mathcal{D}^{\gamma,\eta}}$, we focus on the bound of

$$[\ominus]_r := \big(\mathcal{J}_r^{\mathsf{M}}(x)\boldsymbol{f}(x) + (\mathcal{N}_r^{\mathsf{M}}\boldsymbol{f})(x)\big)_{X^n}$$

for any $|n|_{\mathfrak{s}} < \gamma+2$. If $r \leq \omega(x)^4$, we estimate the $\mathcal{J}^{\mathsf{M}}$ and $\mathcal{N}^{\mathsf{M}}$ terms separately. By using the bound (4.6) for $\mathcal{N}^{\mathsf{M}}$, we have

$$[\ominus]_r \lesssim \sum_{\beta\in(|n|_{\mathfrak{s}}-2,\gamma)} \omega(x)^{(\eta-\beta)\wedge 0} r^{\frac{\beta-|n|_{\mathfrak{s}}-2}{4}} + \omega(x)^{\eta\wedge\beta_0-\gamma} r^{\frac{\gamma-|n|_{\mathfrak{s}}-2}{4}}.$$

Its integration over $r \in [0,\omega(x)^4]$ has an upper bound $\omega(x)^{\eta\wedge\beta_0+2-|n|_{\mathfrak{s}}}$. If $r > \omega(x)^4$, we decompose

$$[\ominus]_r = \frac{1}{n!}\langle \mathbf{R}^{\mathsf{M}}\boldsymbol{f}, \partial_x^n q_r(x,\cdot)\rangle - \sum_{\beta<\gamma,\,|n|_{\mathfrak{s}}\geq\beta+2} \frac{1}{n!}\big\langle \Pi_x^{\mathsf{g}}\big(\boldsymbol{f}(x)\big)_\beta, \partial_x^n q_r(x,\cdot)\big\rangle.$$

Note that, in the sum over β, the term associated with β such that $|n|_{\mathfrak{s}} = \beta+2$ vanishes because of (3.3). Since $\mathbf{R}^{\mathsf{M}}\boldsymbol{f} \in \mathcal{C}^{\eta\wedge\beta_0}$, we have

$$[\ominus]_r \lesssim r^{\frac{\eta\wedge\beta_0-|n|_{\mathfrak{s}}-2}{4}} + \sum_{\beta<\gamma,\,|n|_{\mathfrak{s}}>\beta+2} \omega(x)^{(\eta-\beta)\wedge 0} r^{\frac{\beta-|n|_{\mathfrak{s}}-2}{4}}.$$

Its integration over $r \in [0,\omega(x)^4]$ has an upper bound $\omega(x)^{(\eta\wedge\beta_0+2-|n|_{\mathfrak{s}})\wedge 0}$.

As for the bound of $\|\mathcal{K}^{\mathsf{M}} \boldsymbol{f}\|_{\mathcal{D}^{\gamma,\eta}}$, we focus on the bound of $(\ominus)_r$ similarly to the proof of Theorem 17. For the integrations over $r \le d(x,y)^4$ and $d(x,y)^4 \le r \le \omega(x,y)^4$, we use the same $(\ast)$-decomposition and the $(\bigstar)$-decomposition, respectively. Then, we have the same bounds except the existence of the factor $\omega(x,y)^{\eta\wedge\beta_0-\gamma}$. For the integration over $r \ge \omega(x,y)^4$, we further decompose $(\bigstar)_r^1$ into

$$(\bigstar)_r^1 = \frac{1}{n!}\big\langle \mathbf{R}^{\mathsf{M}} \boldsymbol{f}, (\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}\big\rangle - \frac{1}{n!}\big\langle \Pi_x^{\mathsf{g}} \boldsymbol{f}(x), (\partial^n q_r)_{y,x}^{\gamma+2-|n|_{\mathfrak{s}}}\big\rangle.$$

For the first term, since $\mathbf{R}^{\mathsf{M}} \boldsymbol{f} \in \mathcal{C}^{\eta\wedge\beta_0}$, by using the same remainder formula (3.16), we have the upper bound $\sum_{|\ell|_{\mathfrak{s}}>\gamma+2-|n|_{\mathfrak{s}}} d(x,y)^{|\ell|_{\mathfrak{s}}} r^{\frac{\eta\wedge\beta_0-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}-2}{4}}$. Its integration over $r \in [\omega(x,y)^4, 1]$ has an upper bound

$$\sum_{|\ell|_{\mathfrak{s}}>\gamma+2-|n|_{\mathfrak{s}}} d(x,y)^{|\ell|_{\mathfrak{s}}}\omega(x,y)^{\eta\wedge\beta_0-|n|_{\mathfrak{s}}-|\ell|_{\mathfrak{s}}+2} \lesssim d(x,y)^{\gamma+2-|n|_{\mathfrak{s}}}\omega(x,y)^{\eta\wedge\beta_0-\gamma}$$

since $d(x,y) \le \omega(x,y)$. We also have a similar bound for the remaining terms. ∎

C.2. Proof of Propositions 31 and 32

We provide a sketch of the proof of Propositions 31 and 32, following [21, 41]. Recall that the basis $\mathcal{V}$ of V is the set of all rooted trees with node types $\mathfrak{T}_{\mathrm{n}}$ and edge types $\mathfrak{T}_{\mathrm{e}}$, and with decorations $\mathfrak{n} : N_\tau \to \mathbb{N}^{d+1}$ and $\mathfrak{e} : E_\tau \to \mathbb{N}^{d+1}$. Write $\mathsf{N} = \mathfrak{T}_{\mathrm{n}} \times \mathbb{N}^{d+1}$ and $\mathsf{E} = \mathfrak{T}_{\mathrm{e}} \times \mathbb{N}^{d+1}$. Moreover, the basis $\mathcal{B}$ of U is a subset satisfying Assumption D. For simplicity, we prove the following proposition for V and U, instead of V^* and U^*.

Proposition. (a) The space V (or U) with the operators $\{\overset{\mathsf{e}}{\curvearrowright}\}_{\mathsf{e}\in E}$ (or $\{\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}\}_{\mathsf{e}\in E}$) is an E-multi-pre-Lie algebra.

(b) Let $(W, \{\rhd_{\mathsf{e}}\}_{\mathsf{e}\in\mathsf{E}})$ be an E-multi-pre-Lie algebra, and let

$$\{\varphi_{\mathfrak{b}^n}\}_{(\mathfrak{b},n)\in\mathsf{N}} \ (\text{or } \{\varphi_{\mathfrak{b}^n}\}_{(\mathfrak{b},n)\in\mathsf{N}\cap\mathcal{B}}) \subset W.$$

Then, there exists a unique E-multi-pre-Lie morphism $\varphi : V$ (or U) $\to W$ such that $\varphi(\mathfrak{b}^n) = \varphi_{\mathfrak{b}^n}$ for any $\mathfrak{b}^n \in \mathsf{N}$ (or $\mathsf{N} \cap \mathcal{B}$).

The E-multi-pre-Lie property of $(V, \{\overset{\mathsf{e}}{\curvearrowright}\}_{\mathsf{e}\in E})$

$$(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}) \overset{\mathsf{e}'}{\curvearrowright} \boldsymbol{\eta} - \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} (\boldsymbol{\sigma} \overset{\mathsf{e}'}{\curvearrowright} \boldsymbol{\eta}) = (\boldsymbol{\sigma} \overset{\mathsf{e}'}{\curvearrowright} \boldsymbol{\tau}) \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\eta} - \boldsymbol{\sigma} \overset{\mathsf{e}'}{\curvearrowright} (\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\eta})$$

is proved in a similar way to [41, Proposition 2.2] and [21, Corollary 2.9], so we omit the proof. The same property for $(U, \{\overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}}\}_{\mathsf{e}\in E})$ follows from it. Indeed, by Assumption D1 for the canonical projection $\pi_U : V \to U$, we have

$$(\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}} \boldsymbol{\sigma}) \overset{\mathsf{e}'}{\curvearrowright}_{\mathfrak{b}} \boldsymbol{\eta} - \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright}_{\mathfrak{b}} (\boldsymbol{\sigma} \overset{\mathsf{e}'}{\curvearrowright}_{\mathfrak{b}} \boldsymbol{\eta}) = \pi_U\big((\boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} \boldsymbol{\sigma}) \overset{\mathsf{e}'}{\curvearrowright} \boldsymbol{\eta} - \boldsymbol{\tau} \overset{\mathsf{e}}{\curvearrowright} (\boldsymbol{\sigma} \overset{\mathsf{e}'}{\curvearrowright} \boldsymbol{\eta})\big).$$

In order to prove (b), we introduce the Guin–Oudom extension of the multi-pre-Lie structure. The following is the content of [41, Section 2.2] and [21, Section 3.2].

Definition. Let $(W, \{\rhd_{\mathsf{e}}\}_{\mathsf{e}\in\mathsf{E}})$ be an E-multi-pre-Lie algebra. Let $(W, \mathsf{e}) = \{(a, \mathsf{e})\}_{a\in W}$ be a copy of the linear space W, and denote

$$W^{\oplus\mathsf{E}} := \bigoplus_{\mathsf{e}\in E}(W, \mathsf{e}).$$

Moreover, let $S(W^{\oplus\mathsf{E}})$ be the symmetric algebra of $W^{\oplus\mathsf{E}}$, with unit $\mathbf{1}$. Then, one can define the following linear maps.

- Define the linear map $\rhd_{\mathsf{e}} : W \otimes S(W^{\oplus\mathsf{E}}) \to S(W^{\oplus\mathsf{E}})$ inductively as follows:

$$c \rhd_{\mathsf{e}} \mathbf{1} = 0,$$
$$c \rhd_{\mathsf{e}} \prod_{i=1}^{N}(b_i, \mathsf{e}_i) = \sum_{i=1}^{N}(c \rhd_{\mathsf{e}} b_i, \mathsf{e}_i) \prod_{j\neq i}(b_j, \mathsf{e}_j),$$

  where $c, b_1, \ldots, b_N \in W$ and $\mathsf{e}_1, \ldots, \mathsf{e}_N \in \mathsf{E}$.

- Define the linear map $\rhd : S(W^{\oplus\mathsf{E}}) \otimes W \to W$ inductively as follows:

$$\mathbf{1} \rhd b = b,$$
$$(c, \mathsf{e}) \rhd b = c \rhd_{\mathsf{e}} b,$$
$$\prod_{i=1}^{N}(c_i, \mathsf{e}_i) \rhd b = c_1 \rhd_{\mathsf{e}_1} \left(\prod_{i=2}^{N}(c_i, \mathsf{e}_i) \rhd b\right) - \left(c_1 \rhd_{\mathsf{e}_1} \prod_{i=2}^{N}(c_i, \mathsf{e}_i)\right) \rhd b,$$

  where $c, c_1, \ldots, c_N, b \in W$ and $\mathsf{e}, \mathsf{e}_1, \ldots, \mathsf{e}_N \in \mathsf{E}$. (The last quantity is invariant under the permutations of $(c_1, \mathsf{e}_1), \ldots, (c_N, \mathsf{e}_N)$ because of the multi-pre-Lie property of W, so the extension $\rhd$ is well defined.)

The above extensions keep the multi-pre-Lie morphism property. Indeed, if $\varphi : V \to W$ is an E-multi-pre-Lie morphism, then defining the extension $\varphi : S(V^{\oplus\mathsf{E}}) \to S(W^{\oplus\mathsf{E}})$ by $\varphi((c, \mathsf{e})) := (\varphi(c), \mathsf{e})$ for any $(c, \mathsf{e}) \in (V, \mathsf{E})$, and denoting by $\curvearrowright$ the extension $\curvearrowright$: $S(V^{\oplus\mathsf{E}}) \otimes V \to V$, one has

$$\varphi(X \curvearrowright b) = \varphi(X) \rhd \varphi(b)$$

for any $X \in S(V^{\oplus\mathsf{E}})$ and $b \in V$.

Since $\mathfrak{T}_{\mathsf{e}} = \{\mathcal{I}\}$, we identify E with $\mathbb{N}^{d+1}$. The following formula can be proved by a similar argument to [41, Lemma 2.6] by taking the "Taylor deformation" map Θ introduced in [21] (see Theorem 2.7 and Proposition 3.8 of therein) into account.

Lemma 54. *For any* $(\mathsf{t}, n) \in \mathsf{N} \cap \mathcal{B}$, $\boldsymbol{\tau}_1, \ldots, \boldsymbol{\tau}_a \in \mathcal{B}$, *and* $p_1, \ldots, p_a \in \mathbb{N}^{d+1}$, *one has*

$$\pi_U\left(\mathfrak{b}^n \star \mathop{\bigstar}_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i)\right)$$
$$= \sum_{\substack{q_1\leq p_1,\ldots,q_a\leq p_a\\ q_1+\cdots+q_a\leq n}} (-1)^{|q_1|+\cdots+|q_a|}\binom{n}{q_1,\ldots,q_a}\prod_{i=1}^{a}(\boldsymbol{\tau}_i, p_i - q_i) \curvearrowright_{\mathfrak{b}} \mathfrak{b}^{n-q_1-\cdots-q_a},$$

where $\curvearrowright_{\mathfrak{b}}$ *denotes the extension* $\curvearrowright_{\mathfrak{b}}: S(U^{\oplus \mathsf{E}}) \otimes U \to U$, *and* $\binom{n}{q_1,\dots,q_a}$ *is the multinomial coefficient*

$$\binom{n}{q_1,\dots,q_a} = \frac{n!}{q_1!\cdots q_a!(n-q_1-\cdots-q_a)!}.$$

Then, we can prove the uniqueness part of (b) immediately. Indeed, if $\varphi : U \to W$ is an E-multi-pre-Lie morphism, then it extends to a multi-pre-Lie morphism from $S(U^{\oplus \mathsf{E}})$ to $S(W^{\oplus \mathsf{E}})$ and satisfies

$$\varphi\Bigg(\mathfrak{b}^n \star \bigstar_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i)\Bigg) = \sum_{\substack{q_1\le p_1,\dots,q_a\le p_a\\ q_1+\cdots+q_a\le n}} (-1)^{|q_1|+\cdots+|q_a|}\binom{n}{q_1,\dots,q_a}\prod_{i=1}^{a}(\varphi(\boldsymbol{\tau}_i), p_i-q_i) \rhd \varphi(\mathfrak{b}^{n-q_1-\cdots-q_a})$$

for any $\mathfrak{b}^n \star \bigstar_{i=1}^{a} \mathcal{I}_{p_i}(\boldsymbol{\tau}_i) \in \mathcal{B}$. The right-hand side provides the recursive definition of the map φ, so we can conclude that φ is determined by the values $\varphi(\mathfrak{b}^n)$ for any $\mathfrak{b}^n \in \mathsf{N}$. On the other hand, given $\{\varphi(\mathfrak{b}^n)\}_{(\mathfrak{t},n)\in\mathsf{N}}$, we can prove that the map φ defined by the above formula satisfies indeed the multi-pre-Lie property. See [21, Proposition 2.5 and Corollary 2.10] for details. We do not provide the details here because only the uniqueness part of (b) is used in this paper, especially in Corollary 41.

C.3. Proof of Lemma 46

C.3.1. Reduced coproducts. First, we consider trees with $\mathfrak{n}$- and $\mathfrak{e}$-decorations, without $\mathfrak{o}$-decoration. Recall that SC is a set of strongly conforming trees and C is a set of conforming trees. Set

$$^\circ T := \operatorname{span}(SC), \quad {}^\circ\mathsf{T}^+ := \operatorname{span}(C), \quad {}^\circ\mathsf{U}^- := \mathbb{R}[SC].$$

Definition. We define the following splitting maps.

(1) The linear map $^\circ\mathsf{D} : {}^\circ T \to {}^\circ T \otimes {}^\circ\mathsf{T}^+$ is defined for $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in SC$ by

$$^\circ\mathsf{D}(\tau_{\mathfrak{e}}^{\mathfrak{n}}) := \sum_{\mu\in \mathrm{ST}(\tau)}\sum_{\mathfrak{n}_\mu,\mathfrak{e}'_{\partial\mu}} \frac{1}{\mathfrak{e}'_{\partial\mu}!}\binom{\mathfrak{n}}{\mathfrak{n}_\mu}\mu_{\mathfrak{e}}^{\mathfrak{n}_\mu+\pi\mathfrak{e}'_{\partial\mu}} \otimes \big(\tau/^{\mathrm{blue}}\mu\big)_{\mathfrak{e}+\mathfrak{e}'_{\partial\mu}}^{[\mathfrak{n}-\mathfrak{n}_\mu]_\mu},$$

where $\mathrm{ST}(\tau)$ is the set of all *subtrees* μ of τ which contain the root of τ, and the second sum is over functions $\mathfrak{n} : N_\mu \to \mathbb{N}\times\mathbb{N}^d$ with $\mathfrak{n}_\mu \le \mathfrak{n}$ and functions $\mathfrak{e}'_{\partial\mu} : \partial\mu \to \mathbb{N}\times\mathbb{N}^d$. The algebra morphism

$$^\circ\mathsf{D}^+ : {}^\circ\mathsf{T}^+ \to {}^\circ\mathsf{T}^+ \otimes {}^\circ\mathsf{T}^+$$

is defined by the same formula for $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in C$.

(2) The algebra morphism

$$^{\circ}\mathsf{D}^{-} : {}^{\circ}\mathsf{U}^{-} \to {}^{\circ}\mathsf{U}^{-} \otimes {}^{\circ}\mathsf{U}^{-}$$

is defined by ${}^{\circ}\mathsf{D}^{-}\mathbf{1}_{-} = \mathbf{1}_{-} \otimes \mathbf{1}_{-}$, and for $\tau_{\mathfrak{e}}^{\mathfrak{n}} \in SC$,

$$^{\circ}\mathsf{D}^{-}(\tau_{e}^{\mathfrak{n}}) := \sum_{\varphi \in \mathrm{SF}(\tau)} \sum_{\mathfrak{n}_{\varphi}, \mathfrak{e}'_{\partial\varphi}} \frac{1}{\mathfrak{e}'_{\partial\varphi}!} \binom{\mathfrak{n}}{\mathfrak{n}_{\varphi}} \varphi_{\mathfrak{e}}^{\mathfrak{n}_{\varphi}+\pi \mathfrak{e}'_{\partial\varphi}} \otimes \big(\tau/^{\mathrm{red}}\varphi\big)_{\mathfrak{e}+\mathfrak{e}'_{\partial\varphi}}^{[\mathfrak{n}-\mathfrak{n}]_{\varphi}},$$

where $\mathrm{SF}(\tau)$ is the set of all *subforests* φ of τ *which contain all red nodes of* τ, and the sum over $\mathfrak{n}_{\varphi}$ and $\mathfrak{e}'_{\partial\varphi}$ is taken as in item 1.

(3) The algebra morphism

$$^{\circ}\overline{\mathsf{D}}^{-} : {}^{\circ}\mathsf{T}^{+} \to {}^{\circ}\mathsf{U}^{-} \otimes {}^{\circ}\mathsf{T}^{+}$$

is defined by the same formula as ${}^{\circ}\mathsf{D}^{-}$, but the first sum is restricted to the set $\overline{\mathrm{SF}}(\tau)$ of all subforests $\varphi \in \mathrm{SF}(\tau)$ *which is disjoint with the root of* τ.

Our aim is to show the co-associativities of ${}^{\circ}\mathsf{D}^{\pm}$ and ${}^{\circ}\mathsf{D}$ and ${}^{\circ}\overline{\mathsf{D}}^{-}$. To avoid a confusing calculation, we separate the coproducts into graph part and decoration part. Define simpler coproducts acting on undecorated trees by

$$^{*}\mathsf{D}^{+}\tau := \sum_{\sigma \in \mathrm{ST}(\tau)} \sigma \otimes \big(\tau/^{\mathrm{blue}}\sigma\big),$$

$$^{*}\mathsf{D}^{-}\tau := \sum_{\varphi \in \mathrm{SF}(\tau)} \varphi \otimes \big(\tau/^{\mathrm{red}}\varphi\big),$$

$$^{*}\overline{\mathsf{D}}^{-}\tau := \sum_{\varphi \in \overline{\mathrm{SF}}(\tau)} \varphi \otimes \big(\tau/^{\mathrm{red}}\varphi\big).$$

Given an undecorated tree τ, denote by $\mathbb{X}_{(n,k)}$ the map adding to the node $n \in N_{\tau}$ the $\mathfrak{n}$-decoration $k \in \mathbb{N} \times \mathbb{N}^{d}$, and denote by $\mathbb{I}_{(e,\ell)}$ the map giving to the edge $e \in E_{\tau}$ the $\mathfrak{e}$-decoration $\ell \in \mathbb{N} \times \mathbb{N}^{d}$. Then, any decorated tree $\tau_{\mathfrak{e}}^{\mathfrak{n}}$ is of the form

$$\mathbb{F}\tau = \mathbb{F}_{1} \cdots \mathbb{F}_{N}\tau, \tag{C.7}$$

where τ is an undecorated tree, and $\mathbb{F}_{1}, \dots, \mathbb{F}_{N}$ are family of $\mathbb{X}$-type or $\mathbb{I}$-type operators, applying to pairwise different nodes or edges. Moreover, we define the coproducts of such operators by

$$\mathbb{D}\mathbb{X}_{(n,k)} = \sum_{k' \le k} \binom{k}{k'} \mathbb{X}_{(n,k')} \otimes \mathbb{X}_{(n,k-k')},$$

$$\mathbb{D}\mathbb{I}_{(e,\ell)} = \mathbb{I}_{(e,\ell)} \otimes \mathbb{X}_{(e_{-},0)} + \sum_{\ell'} \frac{1}{\ell'!} \mathbb{X}_{(e_{-},\ell')} \otimes \mathbb{I}_{(e,\ell+\ell')},$$

where e_- denotes the node from where the edge e leaves. For the products of pairwise disjoint such operators, define

$$\mathbb{D}\mathbb{F} := (\mathbb{D}\mathbb{F}_1)\cdots(\mathbb{D}\mathbb{F}_N). \tag{C.8}$$

There are two remarks about this identity. First, the right-hand side does not depend on the order of $\mathbb{F}_1, \dots, \mathbb{F}_N$. To prove it, we have only to show that $\mathbb{D}\mathbb{X}_{(e_-,k)}\mathbb{D}\mathbb{I}_{(e,\ell)} = \mathbb{D}\mathbb{I}_{(e,\ell)}\mathbb{D}\mathbb{X}_{(e_-,k)}$ and $\mathbb{D}\mathbb{I}_{(e,k)}\mathbb{D}\mathbb{I}_{(f,\ell)} = \mathbb{D}\mathbb{I}_{(f,\ell)}\mathbb{D}\mathbb{I}_{(e,k)}$ for e, f such that $e_- = f_-$, because for other pairs of operators $\mathbb{F}$ and $\mathbb{G}$, the operators $\mathbb{D}\mathbb{F}$ and $\mathbb{D}\mathbb{G}$ apply to different nodes or edges. These identities can be checked easily by definitions. Second, (C.8) holds even if there are distinct i and j such that $\mathbb{F}_i$ and $\mathbb{F}_j$ apply to the same *node*. Let $\mathbb{F}_i = \mathbb{X}_{(n,k)}$ and $\mathbb{F}_j = \mathbb{X}_{(n,\ell)}$. Since $\mathbb{X}_{(n,k)}\mathbb{X}_{(n,\ell)} = \mathbb{X}_{(n,k+\ell)}$, we have

$$\mathbb{D}(\mathbb{X}_{(n,k)}\mathbb{X}_{(n,\ell)}) = \sum_{m\le k+\ell} \binom{k+\ell}{m} \mathbb{X}_{(n,m)} \otimes \mathbb{X}_{(n,k+\ell-m)}.$$

On the other hand, we have

$$(\mathbb{D}\mathbb{X}_{(n,k)})(\mathbb{D}\mathbb{X}_{(n,\ell)}) = \sum_{a\le k,\, b\le \ell} \binom{k}{a}\binom{\ell}{b} \mathbb{X}_{(n,a+b)} \otimes \mathbb{X}_{(n,k+\ell-a-b)}.$$

We can see that the right-hand sides coincide by using Chu–Vandermonde identity

$$\sum_{a+b=m} \binom{k}{a}\binom{\ell}{b} = \binom{k+\ell}{m}.$$

Therefore, we have $(\mathbb{D}\mathbb{X}_{(n,k)})(\mathbb{D}\mathbb{X}_{(n,\ell)}) = \mathbb{D}\mathbb{X}_{(n,k+\ell)} = (\mathbb{D}\mathbb{X}_{(n,\ell)})(\mathbb{D}\mathbb{X}_{(n,k)})$.

At this stage, we see that the coproducts ${}^\circ\mathsf{D}^{(\cdot,+,-)}$ apply to the decorated tree (C.7) by the forms

$${}^\circ\mathsf{D}^{(\cdot,+)}(\mathbb{F}\tau) = (\mathbb{D}\mathbb{F})({}^*\mathsf{D}^+\tau), \quad {}^\circ\mathsf{D}^-(\mathbb{F}\tau) = (\mathbb{D}\mathbb{F})({}^*\mathsf{D}^-\tau). \tag{C.9}$$

On the right-hand side of (C.9), be careful that $\mathbb{D}\mathbb{F}$ acts on subtrees and contracted trees. For an $\mathbb{X}$-type operator, if $n \notin N_\sigma$, then set $\mathbb{X}_{(n,k)}\sigma = \mathbf{1}_{k=0}\sigma$. On a contracted tree τ/σ, the $\mathbb{X}$-type operator acts of the form $\mathbb{X}_{([n],k)}$, where $[n]$ denotes the equivalence class in the contraction $\tau \to \tau/\sigma$. Hence,

$$(\mathbb{D}\mathbb{X}_{(n,k)})(\sigma \otimes (\tau/\sigma)) = \begin{cases} \sum_{k'\le k} \binom{k}{k'} \mathbb{X}_{(n,k')}\sigma \otimes \mathbb{X}_{([n],k-k')}(\tau/\sigma), & n \in N_\sigma, \\ \sigma \otimes \mathbb{X}_{([n],k)}(\tau/\sigma), & n \notin N_\sigma. \end{cases}$$

For an $\mathbb{I}$-type operator, if $e \notin E_\sigma$ (resp., $e \notin E_{\tau/\sigma}$), then set $\mathbb{I}_{(e,\ell)}\sigma = 0$ (resp., $\mathbb{I}_{(e,\ell)}(\tau/\sigma) = 0$). Combining with the definition of $\mathbb{X}_{(n,k)}$, we have

$$\begin{aligned}&(\mathbb{D}\mathbb{I}_{(e,\ell)})(\sigma \otimes (\tau/\sigma)) \\ &\quad = \begin{cases} \mathbb{I}_{(e,\ell)}\sigma \otimes (\tau/\sigma), & \text{for } e \in E_\sigma, \\ \sum_{\ell'} \frac{1}{\ell'!} \mathbb{X}_{(e_-,\ell')}\sigma \otimes \mathbb{I}_{(e,\ell+\ell')}(\tau/\sigma), & \text{for } e \in \partial\sigma, \\ \sigma \otimes \mathbb{I}_{(e,\ell)}(\tau/\sigma), & \text{for } e \in E_\tau \setminus (E_\sigma \cup \partial\sigma). \end{cases}\end{aligned}$$

These conventions show that the identities (C.9) hold.

C.3.2. Co-associativity.

Lemma 55. *One has the co-associativity formulas*

$$\big({}^\circ\mathsf{D}\otimes\mathrm{Id}\big){}^\circ\mathsf{D}=\big(\mathrm{Id}\otimes{}^\circ\mathsf{D}^+\big){}^\circ\mathsf{D},\qquad \big({}^\circ\mathsf{D}^+\otimes\mathrm{Id}\big){}^\circ\mathsf{D}^+=\big(\mathrm{Id}\otimes{}^\circ\mathsf{D}^+\big){}^\circ\mathsf{D}^+,$$
$$\big({}^\circ\mathsf{D}^-\otimes\mathrm{Id}\big){}^\circ\mathsf{D}^-=\big(\mathrm{Id}\otimes{}^\circ\mathsf{D}^-\big){}^\circ\mathsf{D}^-,\qquad \big({}^\circ\mathsf{D}^-\otimes\mathrm{Id}\big){}^\circ\overline{\mathsf{D}}^-=\big(\mathrm{Id}\otimes{}^\circ\overline{\mathsf{D}}^-\big){}^\circ\overline{\mathsf{D}}^-.$$

Proof. We prove the identity

$$({}^\circ\mathsf{D}\otimes\mathrm{Id}){}^\circ\mathsf{D}=(\mathrm{Id}\otimes{}^\circ\mathsf{D}^+){}^\circ\mathsf{D};\tag{C.10}$$

the other identities are proved similarly. By the commutation relation (C.9), we have

$$\big(\mathrm{Id}\otimes{}^\circ\mathsf{D}^+\big){}^\circ\mathsf{D}\,\mathbb{F}\tau=\big(\mathrm{Id}\otimes{}^\circ\mathsf{D}^+\big)(\mathbb{D}\mathbb{F})({}^*\mathsf{D}^+\tau)=\big((\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}\big)\big(\mathrm{Id}\otimes{}^*\mathsf{D}^+\big){}^*\mathsf{D}^+\tau.$$

Hence, it is sufficient for proving (C.10) to show the two identities

$$\big({}^*\mathsf{D}^+\otimes\mathrm{Id}\big){}^*\mathsf{D}^+=\big(\mathrm{Id}\otimes{}^*\mathsf{D}^+\big){}^*\mathsf{D}^+,\tag{C.11}$$
$$(\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}=(\mathbb{D}\otimes\mathrm{Id})\mathbb{D}\mathbb{F}.\tag{C.12}$$

It is not difficult to show (C.11) by the definition of $\mathrm{ST}(\tau)$, by noting that

$$\mathrm{ST}(\tau/\sigma)=\big\{\eta/\sigma;\ \sigma\subset\eta\subset\tau\big\}$$

and $(\tau/\sigma)/(\eta/\sigma)=\tau/\eta$. Next, we show (C.12). By the multiplicativity (C.8), for any family $\mathbb{F}_1,\dots,\mathbb{F}_N$ of $\mathbb{X}$-type or $\mathbb{I}$-type operators, applying to pairwise different nodes or edges, one has

$$\begin{aligned}(\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}&=(\mathrm{Id}\otimes\mathbb{D})\big((\mathbb{D}\mathbb{F}_1)\cdots(\mathbb{D}\mathbb{F}_N)\big)\\&=\big((\mathrm{Id}\otimes\mathbb{D})(\mathbb{D}\mathbb{F}_1)\big)\cdots\big((\mathrm{Id}\otimes\mathbb{D})(\mathbb{D}\mathbb{F}_N)\big).\end{aligned}$$

In the last equality, we use the fact that (C.8) holds even if multiple operators act on the same node. If (C.12) holds for $\mathbb{F}_1,\dots,\mathbb{F}_N$, then

$$\begin{aligned}(\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}&=\big((\mathbb{D}\otimes\mathrm{Id})(\mathbb{D}\mathbb{F}_1)\big)\cdots\big((\mathbb{D}\otimes\mathrm{Id})(\mathbb{D}\mathbb{F}_N)\big)\\&=(\mathbb{D}\otimes\mathrm{Id})\big((\mathbb{D}\mathbb{F}_1)\cdots(\mathbb{D}\mathbb{F}_N)\big)=(\mathbb{D}\otimes\mathrm{Id})\mathbb{D}\mathbb{F}.\end{aligned}$$

Therefore, it is sufficient to show (C.12) for $\mathbb{F}=\mathbb{X}_{(n,k)}$ and $\mathbb{I}_{(e,\ell)}$. For $\mathbb{F}=\mathbb{X}_{(n,k)}$, we have

$$\begin{aligned}(\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{X}_{(n,k)}&=\sum_{k'\le k}\binom{k}{k'}\mathbb{X}_{(n,k')}\otimes\mathbb{D}\mathbb{X}_{(n,k-k')}\\&=\sum_{k'\le k,\,k''\le k-k'}\binom{k}{k'}\binom{k-k'}{k''}\mathbb{X}_{(n,k')}\otimes\mathbb{X}_{(n,k'')}\otimes\mathbb{X}_{(n,k-k'-k'')}\\&=\sum_{k',k'';k'+k''\le n}\binom{k}{k',k''}\mathbb{X}_{(n,k')}\otimes\mathbb{X}_{(n,k'')}\otimes\mathbb{X}_{(n,k-k'-k'')}.\end{aligned}$$

We have the same expansion from $(\mathbb{D}\otimes\mathrm{Id})\mathbb{D}\mathbb{X}_{(n,k)}$, so (C.12) holds for $\mathbb{X}$-type operators. We can prove the same result for $\mathbb{I}$-type operators by similar computations. ∎

Now, we consider the extended decoration.

Lemma 56. *One has the co-associativity formulas*

$$(\mathsf{D}\otimes \mathrm{Id})\mathsf{D} = \big(\mathrm{Id}\otimes \mathsf{D}^{+}\big)\mathsf{D}, \qquad \big(\mathsf{D}^{+}\otimes \mathrm{Id}\big)\mathsf{D}^{+} = \big(\mathrm{Id}\otimes \mathsf{D}^{+}\big)\mathsf{D}^{+},$$
$$(\mathsf{D}^{-}\otimes \mathrm{Id})\mathsf{D}^{-} = \big(\mathrm{Id}\otimes \mathsf{D}^{-}\big)\mathsf{D}^{-}, \quad \big(\mathsf{D}^{-}\otimes \mathrm{Id}\big)\bar{\mathsf{D}}^{-} = \big(\mathrm{Id}\otimes \bar{\mathsf{D}}^{-}\big)\bar{\mathsf{D}}^{-}.$$

Proof. We consider the first and third identities; the other identities are proved similarly. In this proof, denote by $\overline{\tau} = \tau_{\mathfrak{e}}^{\mathfrak{n}}$ a generic decorated tree without $\mathfrak{o}$-decoration, and write $\overline{\tau}^{\mathfrak{o}}$ for $\tau_{\mathfrak{e}}^{\mathfrak{n},\mathfrak{o}}$. As in Sections 2.2 and 5.1, we use a shorthand notation

$${}^{\circ}\mathsf{D}\overline{\tau} = \sum_{\overline{\sigma}\leq\overline{\tau}} \overline{\sigma}\otimes(\overline{\tau}/\overline{\sigma}), \quad {}^{\circ}\mathsf{D}^{-}\overline{\tau} = \sum_{\overline{\varphi}\trianglelefteq\overline{\tau}} \overline{\varphi}\otimes(\overline{\tau}/\overline{\varphi}).$$

Then, we can write

$$\mathsf{D}\overline{\tau}^{\mathfrak{o}} = \sum_{\overline{\sigma}} \overline{\sigma}^{\mathfrak{o}}\otimes(\overline{\tau}/\overline{\sigma})^{\mathfrak{o}|_{\tau\setminus\sigma}}, \quad \mathsf{D}^{-}\overline{\tau}^{\mathfrak{o}} = \sum_{\overline{\varphi}} \overline{\varphi}^{\mathfrak{o}}\otimes(\overline{\tau}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\varphi})}.$$

Recall that $\mathfrak{o}(\overline{\varphi}) : N_{\tau/\varphi} \to \mathbb{Z}[\beta_0]$ is a function giving the value $|\overline{\tau}_j|$, where $\overline{\tau}_j$ is a connected component if φ, to the node $[\tau_j] \in N_{\tau/\varphi}$. We obtain the co-associativity of D from the co-associativity of ${}^{\circ}\mathsf{D}$, noting that

$$(\overline{\tau}/\overline{\sigma})/(\overline{\eta}/\overline{\sigma}) = \overline{\tau}/\overline{\eta}, \quad \mathfrak{o}|_{(\tau\setminus\sigma)\setminus(\eta\setminus\sigma)} = \mathfrak{o}|_{\tau\setminus\eta}$$

for any $\overline{\eta}\leq\overline{\sigma}\leq\overline{\tau}$. To prove the co-associativity of D^{-}, noting that

$$(\mathsf{D}^{-}\otimes\mathrm{Id})\mathsf{D}^{-}\overline{\tau}^{\mathfrak{o}} = \sum_{\overline{\psi}\trianglelefteq\overline{\varphi}\trianglelefteq\overline{\tau}} \overline{\psi}^{\mathfrak{o}}\otimes(\overline{\varphi}/\overline{\psi})^{\mathfrak{o}+\mathfrak{o}(\overline{\psi})}\otimes(\overline{\tau}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\varphi})}$$

and

$$\begin{aligned}(\mathrm{Id}\otimes\mathsf{D}^{-})\mathsf{D}^{-}\overline{\tau}^{\mathfrak{o}} &= \sum_{\overline{\psi}\trianglelefteq\overline{\tau}} \overline{\psi}^{\mathfrak{o}}\otimes\mathsf{D}^{-}(\overline{\tau}/\overline{\psi})^{\mathfrak{o}+\mathfrak{o}(\overline{\psi})}\\ &= \sum_{\overline{\psi}\trianglelefteq\overline{\varphi}\trianglelefteq\overline{\tau}} \overline{\psi}^{\mathfrak{o}}\otimes(\overline{\varphi}/\overline{\psi})^{\mathfrak{o}+\mathfrak{o}(\overline{\psi})}\otimes(\overline{\tau}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\psi})+\mathfrak{o}(\overline{\varphi}/\overline{\psi})},\end{aligned}$$

it is sufficient to show that $\mathfrak{o}(\overline{\varphi}) = \mathfrak{o}(\overline{\psi}) + \mathfrak{o}(\overline{\varphi}/\overline{\psi})$ as a function on $N_{\tau/\varphi}$. This holds true because $|\overline{\varphi}|' = |\overline{\psi}|' + |\overline{\varphi}/\overline{\psi}|'$. ∎

C.3.3. Co-interaction.

Lemma 57. *One has the co-interaction formulas*

$$\begin{aligned}\mathcal{M}^{(13)}\big({}^{\circ}\mathsf{D}^{-}\otimes{}^{\circ}\bar{\mathsf{D}}^{-}\big){}^{\circ}\mathsf{D} &= (\mathrm{Id}\otimes{}^{\circ}\mathsf{D}){}^{\circ}\mathsf{D}^{-},\\ \mathcal{M}^{(13)}\big({}^{\circ}\bar{\mathsf{D}}^{-}\otimes{}^{\circ}\bar{\mathsf{D}}^{-}\big){}^{\circ}\mathsf{D}^{+} &= \big(\mathrm{Id}\otimes{}^{\circ}\mathsf{D}^{+}\big){}^{\circ}\bar{\mathsf{D}}^{-},\end{aligned} \tag{C.13}$$

and

$$\begin{aligned}\mathcal{M}^{(13)}\big(\mathsf{D}^{-}\otimes\bar{\mathsf{D}}^{-}\big)\mathsf{D} &= (\mathrm{Id}\otimes\mathsf{D})\mathsf{D}^{-},\\ \mathcal{M}^{(13)}\big(\bar{\mathsf{D}}^{-}\otimes\bar{\mathsf{D}}^{-}\big)\mathsf{D}^{+} &= (\mathrm{Id}\otimes\mathsf{D}^{+})\bar{\mathsf{D}}^{-}.\end{aligned} \tag{C.14}$$

Proof. Consider the first identity of (C.13) and the first identity of (C.14); the two other identities are proved similarly. By the commutation relations (C.9), identity (C.13) rewrites

$$\mathcal{M}^{(13)}\big((\mathbb{D}\otimes\mathbb{D})\mathbb{D}\mathbb{F}\big)\big({}^*\mathsf{D}^-\otimes{}^*\overline{\mathsf{D}}^-\big){}^*\mathsf{D}^+\tau=\big((\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}\big)(\mathrm{Id}\otimes{}^*\mathsf{D}^+)^*\mathsf{D}^-. \tag{C.15}$$

By the multiplicativity (C.8) of $\mathbb{D}$, it is sufficient to show (C.15) for the operators $\mathbb{F}=\mathbb{X}_{(n,k)}$ and $\mathbb{I}_{(e,\ell)}$. By definition,

$$\big({}^*\mathsf{D}^-\otimes{}^*\overline{\mathsf{D}}^-\big){}^*\mathsf{D}^+\tau=\sum_{\sigma\in\mathrm{ST}(\tau)}\ \sum_{\varphi\in\mathrm{SF}(\sigma),\psi\in\overline{\mathrm{SF}}(\tau/\sigma)}\varphi\otimes(\sigma/^{\mathrm{red}}\varphi)\otimes\psi\otimes(\tau/^{\mathrm{blue}}\sigma)/^{\mathrm{red}}\psi.$$

Note that φ and ψ are disjoint subforests of τ because of the definition of $\overline{\mathrm{SF}}$. Thus, we have

$$\begin{aligned}
&\mathcal{M}^{(13)}\big((\mathbb{D}\otimes\mathbb{D})\mathbb{D}\mathbb{X}_{(n,k)}\big)\big(\varphi\otimes(\sigma/^{\mathrm{red}}\varphi)\otimes\psi\otimes(\tau/^{\mathrm{blue}}\sigma)/^{\mathrm{red}}\psi\big)\\
&\quad=\mathcal{M}^{(13)}\sum_{k=a+b+c+d}\frac{k!}{a!b!c!d!}\mathbb{X}_{(n,a)}\varphi\otimes\mathbb{X}_{(n,b)}(\sigma/^{\mathrm{red}}\varphi)\\
&\qquad\otimes\mathbb{X}_{(n,c)}\psi\otimes\mathbb{X}_{(n,d)}\big((\tau/^{\mathrm{blue}}\sigma)/^{\mathrm{red}}\psi\big)\\
&\quad=\sum_{k=a+b+d}\frac{k!}{a!b!d!}\mathbb{X}_{(n,a)}(\varphi\psi)\otimes\mathbb{X}_{(n,b)}(\sigma/^{\mathrm{red}}\varphi)\otimes\mathbb{X}_{(n,d)}\big((\tau/^{\mathrm{blue}}\sigma)/^{\mathrm{red}}\psi\big)\\
&\quad=\big((\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{X}_{(n,k)}\big)\big(\varphi\psi\otimes(\sigma/^{\mathrm{red}}\varphi)\otimes\big(\tau/^{\mathrm{blue}}\sigma\big)/^{\mathrm{red}}\psi\big),
\end{aligned}$$

since either of a and c has to be 0 in the second line. It is not difficult to show a similar equality for $\mathbb{F}=\mathbb{I}_{(e,\ell)}$. Hence, we have

$$\mathcal{M}^{(13)}\big((\mathbb{D}\otimes\mathbb{D})\mathbb{D}\mathbb{F}\big)\big({}^*\mathsf{D}^-\otimes{}^*\overline{\mathsf{D}}^-\big){}^*\mathsf{D}^+\tau=\big((\mathrm{Id}\otimes\mathbb{D})\mathbb{D}\mathbb{F}\big)\mathcal{M}^{(13)}\big({}^*\mathsf{D}^-\otimes{}^*\overline{\mathsf{D}}^-\big){}^*\mathsf{D}^+\tau.$$

Since it is not difficult to show the co-interaction formula

$$\mathcal{M}^{(13)}\big({}^*\mathsf{D}^-\otimes{}^*\overline{\mathsf{D}}^-\big){}^*\mathsf{D}^+\tau=\big(\mathrm{Id}\otimes{}^*\mathsf{D}^+\big){}^*\mathsf{D}^-,$$

identity (C.15) follows as a consequence.

Next, we consider (C.14). By definition,

$$\begin{aligned}
&\mathcal{M}^{(13)}\big(\mathsf{D}^-\otimes\overline{\mathsf{D}}^-\big)\mathsf{D}\overline{\tau}^{\mathfrak{o}}\\
&\quad=\sum_{\overline{\sigma}\le\overline{\tau},\overline{\varphi}\trianglelefteq\overline{\sigma},\overline{\psi}\trianglelefteq\overline{\tau}/\overline{\sigma}}\overline{\varphi}^{\mathfrak{o}}\,\overline{\psi}^{\mathfrak{o}|_{\tau\setminus\sigma}}\otimes(\overline{\sigma}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\varphi})}\otimes\big((\overline{\tau}/\overline{\sigma})/\overline{\psi}\big)^{\mathfrak{o}|_{\tau\setminus\sigma}+\mathfrak{o}(\overline{\psi})}
\end{aligned}$$

and

$$(\mathrm{Id}\otimes\mathsf{D})\mathsf{D}^-\overline{\tau}^{\mathfrak{o}}=\sum_{\overline{\zeta}\trianglelefteq\overline{\tau},\overline{\eta}\le\overline{\tau}/\overline{\zeta}}\overline{\zeta}^{\mathfrak{o}}\otimes\overline{\eta}^{\mathfrak{o}+\mathfrak{o}(\overline{\zeta})}\otimes\big((\overline{\tau}/\overline{\zeta})/\overline{\eta}\big)^{(\mathfrak{o}+\mathfrak{o}(\overline{\zeta}))|_{(\tau/\zeta)\setminus\eta}}.$$

The cointeraction between $^\circ\mathsf{D}$ and $^\circ\mathsf{D}^-$ implies that the *change of variables*

$$\overline{\zeta}\leftrightarrow\overline{\varphi}\,\overline{\psi},\quad\overline{\eta}\leftrightarrow\overline{\sigma}/\overline{\varphi}$$

is possible. Since σ and ψ are disjoint,

$$(\overline{\varphi}\overline{\psi})^{\mathfrak{o}} = \overline{\varphi}^{\mathfrak{o}}\, \overline{\psi}^{\mathfrak{o}|_{\tau\setminus\sigma}}, \quad (\overline{\sigma}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\varphi}\overline{\psi})} = (\overline{\sigma}/\overline{\varphi})^{\mathfrak{o}+\mathfrak{o}(\overline{\varphi})},$$
$$\big((\overline{\tau}/\overline{\sigma})/\overline{\psi}\big)^{(\mathfrak{o}+\mathfrak{o}(\overline{\varphi}\overline{\psi}))|_{(\tau/(\varphi\psi))\setminus(\sigma/\varphi)}} = \big((\overline{\tau}/\overline{\sigma})/\overline{\psi}\big)^{\mathfrak{o}|_{\tau\setminus\sigma}+\mathfrak{o}(\overline{\psi})}.$$

Thus, (C.14) follows. ∎

D. Comments

Section 1. Regularity structures theory has its roots in T. Lyons' theory of rough paths and rough differential equations [72]. This theory deals with controlled ordinary differential equations

$$dx_t = V(x_t)dh_t,$$

with controls h of low regularity, say, α-Hölder. For $\alpha > 1/2$, Young integration theory allows to make sense of the equation as a fixed-point problem for an integral equation. As one expects a solution path to be α-Hölder, the product $V(x_t)dh_t$ makes sense as a distribution on $\mathbb{R}_+$ iff $\alpha + (\alpha - 1) > 0$, that is, $\alpha > 1/2$. One of Lyons' deep insights was to realize that what really governs the dynamics is not the $\mathbb{R}^\ell$-valued control h, say, but rather a finite collection of its iterated integrals. The latter are ill-defined when $\alpha \leq 1/2$, and a rough path is the a priori datum of quantities playing their role. Natural algebraic and size constraints on these objects are then sufficient to set the entire theory. These constraints are similar to the constraints that define the g-part of a model. Several reformulations of rough paths theory were given after Lyons' seminal work: Davie's numerical scheme approach [33], Gubinelli's controlled paths approach [48, 49], Friz and Victoir's limit ODE picture [44], and Bailleul's approximate flow-to-flow approach [2], amongst others. Gubinelli's versatile notion of controlled paths was a direct source of inspiration for the construction of regularity structures.

Other tools than regularity structures have been developed for the study of singular stochastic PDEs. None of them offers presently a complete alternative to regularity structures.

- Gubinelli, Imkeller, and Perkowski laid in [50] the foundations of paracontrolled calculus, which were developed by Bailleul and Bernicot [3–5]. While the fundamental notions of regularity structures involve pointwise expansions, paracontrolled calculus uses paraproducts as a mean for making sense of what it means to look like a reference quantity. See [51] for lecture notes on the subject and [52] for an overview on the subject, both by Gubinelli and Perkowski.

  In a nutshell, the starting point of the paracontrolled approach to the study of singular stochastic PDEs is the decomposition of a product of two distributions f, g into

$$fg = \big(\mathsf{P}_f g + \mathsf{P}_g f\big) + \Pi(f, g).$$

This decomposition is obtained in a Fourier picture of the product by splitting the convolution into what happens far from the diagonal $(\mathsf{P}_f g + \mathsf{P}_g f)$ from what happens near the diagonal $\Pi(f, g)$. This decomposition isolates in the resonant term $\Pi(f, g)$ what does not make sense in a general product, the paraproduct terms $\mathsf{P}_f g, \mathsf{P}_g f$ being always well defined. The definition of $\mathsf{P}_f g$ allows to think of it as a modulation of g by f and give meaning to what it means for a distribution/function u to look like another distribution/function g

$$u = \mathsf{P}_{u'} g + u^\sharp,$$

for a function u' and a distribution/function $u^\sharp$ that is more regular than g. The role of modelled distributions is played in a paracontrolled setting by systems $(u_a)_{a\in\mathcal{A}}$ of paracontrolled distributions/functions

$$u_a = \sum_{\tau\in\mathcal{T};|a\tau|\leq n\alpha} \mathsf{P}_{u_{a\tau}}[\tau] + u_a^\sharp \tag{D.1}$$

indexed by the set $\mathcal{A}$ of words over an alphabet $\mathcal{T}$, with remainders $u_a^\sharp$ sufficiently regular. The reference distributions/functions $[\tau]$ somehow play the role of $\Pi\tau$ and the $(u_a)_{a\in\mathcal{A}}$ the role of the $(u_\tau)_{\tau\in T,|\tau|<\gamma}$. (Bailleul and Hoshino's work [8, 9] on the relations between paracontrolled calculus and regularity structures make that link clear.) Identity (D.1) is an analogue of the notion of modelled distribution. While the definition of the latter involves pointwise comparisons, here the comparison is somehow done in "momentum space", although not in a pointwise sense. The core point of the paracontrolled analysis of a (system of) singular stochastic PDE(s) is that we end up dealing with ill-defined terms of the form

$$\Pi'\big(u, \zeta^{(i)}\big) \tag{D.2}$$

for operators Π' that have similar properties as the resonant operator Π, and possibly multi-dimensional functionals $\zeta^{(i)}$ of the noise ζ. It turns out that while an expression like (D.2) does not make sense for a generic u, it makes sense on a restricted class of u of the form (D.1) *provided* one can make sense of the terms $\Pi'(\tau, \zeta^{(i)})$. The analysis of a given (system of) singular stochastic PDE(s) gives an inductive definition of the $\zeta^{(i)}$ and the τ's. Compared to the regularity structures setting, the datum of all the $\Pi'(\tau, \zeta^{(i)})$ plays the role of the datum of a model. The inductive/tree structure of the elements of a regularity structure takes here the form of the inductive definition of the τ's and $\zeta^{(i)}$'s. A systematic treatment of renormalization operations within paracontrolled calculus has not been invented yet. The links between the regularity structure and paracontrolled settings detailed in Bailleul and Hoshino's works [8, 9] allow however to transport the renormalization machinery of regularity structures into the setting paracontrolled calculus. What is missing presently is an independent, purely paracontrolled, approach of the renormalization problem.

- Otto and Weber [76] developed jointly with Sauer and Smith [75] a variant of regularity structures that is more in the flavour of rough paths theory. See in also their most recent joint works [69, 70] with Linares, Tempelmayr, and Tsatsoulis to see how far they were able to go. Most concepts and objects from regularity structures have counterparts in their setting. It was specifically designed and used for the analysis of a number of quasilinear singular stochastic PDEs. Some of these equations can be approached using the original first-order paracontrolled calculus as in Bailleul, Debussche, and Hofmanová's work [7] or a variant of it using paracomposition operators, as in Furlan and Gubinelli's work [46]. See also [13, 47] for extensions of paracontrolled calculus and regularity structures designed for the study of a whole class of quasilinear singular stochastic PDEs.
- Kupiainen and Marcozzi managed in [67, 68] to implement a renormalization group approach to the KPZ and Φ^4_3 equations. The starting point of their strategy consists in decomposing the resolution operator $\mathbf{K}^{-1}$ involved in the Picard formulation of the equation as a sum of operators $\mathbf{K}_n^{-1}$ turning distributions into smooth functions that vary essentially only up to scale 2^{-n}. The approximate renormalized dynamics will take the form

$$u_N = \left(\sum_{n=0}^{N} \mathbf{K}_n^{-1}\right)\big(F_N(u_N, \partial u_N; \zeta)\big)$$

in a simplified problem where the initial condition was taken to be null. The noise is left untouched, with no problem for defining the nonlinearity on the right-hand side since u_N is smooth. So, one has

$$u_N = u_N^0 + \cdots + u_N^N,$$

where each term u_N^n is morally varying only up to scale 2^{-n}. The point is now to see that one can choose the nonlinearity F_N in such a way that each u_N^n is the solution of an equation of the form

$$u_N^n = \left(\sum_{m=n}^{N} \mathbf{K}_n^{-1}\right)\big(F_N^n(u_N^n, \partial u_N^n; \zeta)\big)$$

for a nonlinearity F_N^n and is converging in a proper space as N goes to infinity. This is done via the use of rescaling operators, taking profit from the exact scaling property of the heat kernel, by turning the problem of convergence of each u_N^n into the problem of the convergence of the family of rescaled versions of the functions F_N^n – taking advantage of the fact that the former is a continuous function of the latter. The overall convergence of u_N as N goes to ∞ is somehow similar to the well-known fact that a sum of functions $\sum_n u^n$ converges in an α-Hölder space if u^n is localized in Fourier space on a ball of size 2^n and has uniform norm of order $2^{-\alpha n}$.
This approach was improved a lot by some recent works of P. Duch [34–36], who traded the above discrete scale decomposition for a continuous scale decomposition

and uncovered a certain structure on the cumulants of some functionals of the noise that pave the way to the development of a robust approach to some classes of singular equations. One can look at Chandra and Ferdinand's work [25] for an application of this approach to the generalized KPZ equation.

Section 2. The functional setting adopted here draws inspiration from [3–5,76]. The main result of this section is the reconstruction theorem.

Several proofs of the reconstruction theorem are available now, in addition to Hairer's original proof. Gubinelli, Imkeller, and Perkowski gave in [50] an alternative construction of the reconstruction map using a paraproduct-like operator. Singh and Teichmann showed in [77] how it can be understood as the continuous extension of an elementary reconstruction operator defined on a set of smooth modelled distributions. Otto and Weber have an analogue of the reconstruction map in their rough paths-like setting [75, 76]. Caravenna and Zambotti's recent work [23] provide a robust version of the reconstruction theorem in a setting free of any reference to regularity structures. The notion of coherent germ turns it into a particularly versatile tool. See [58, 63, 71] for versions of the reconstruction theorem in functional settings different from Hölder spaces. See also our previous work [8] for a paracontrolled representation of the reconstruction operator that refines over a similar flavoured representation given in [50, Theorem 6.10].

The reconstruction theorem takes its place in the history of a family of statements producing "transcendental" objects, i.e., objects constructed by limiting procedures, from families of objects satisfying constraints involving no limiting procedures. The one-step Euler scheme for solving ordinary differential equations characterizes for instance uniquely their flows under sufficient regularity conditions on the vector fields. In its simplest form, for the equation $\dot{y} = -y$, in $\mathbb{R}$, it yields the elementary identity $(1 - t/n)^n \to e^{-t}$, as n goes to ∞. It takes a more elaborate form in Hille's approximation $((\mathrm{Id} + \frac{t}{n} G)^{-1})^n$ of the semigroup $(e^{-tG})_{t>0}$ generated by an unbounded operator G under well-known conditions. Chernov's theorem [29] on families of strongly continuous perturbations of the identity used for constructing $(e^{-tG})_{t>0}$ has a similar flavour. So is the C^1-approximate flow-to-flow machinery of [2], which provides a far-reaching generalization of Lyons' extension theorem in rough paths theory and Gubinelli and Feyel and de la Pradelle's sewing lemma [37, 38, 48]. All these statements characterize uniquely a transcendental object as the unique object close to a family of objects satisfying a "$O(1)$ condition", involving no limiting procedure. The characterizing identity (2.28) for the reconstruction is of that form when the reconstruction operator is unique. This kind of situation allows to build a calculus for the transcendental objects from an elementary calculus on their generators.

The space of models over a given regularity structure is nonlinear. Bailleul and Hoshino showed in [8, 9] how to parametrize this space by a linear space using the tools of paracontrolled calculus. The set of **K**-admissible models on a given regularity structure turns out in particular to be parametrized by the data for each τ of negative homogeneity of a $|\tau|$-Hölder distribution, describing somehow the most regular part of the distribution

$\Pi\tau$. This has a number of consequences, such as an extension theorem similar to Lyons' extension theorem in rough paths theory.

Proposition 9, giving a definition of the image of a modelled distribution by a nonlinear map, has a counterpart in paracontrolled calculus, generalizing Bony's paralinearisation formula to an arbitrary order – see Section 2 of Bailleul and Bernicot's work [4].

Section 3. The proof of the continuity result for $\mathcal{K}^{\mathsf{M}}$ is an adaptation of the material from Hairer's groundbreaking work [53] to the functional setting adopted here. It is called by Hairer the multilevel Schauder estimates. The construction of **K**-admissible models from Section 3.5 is adapted from Bailleul and Hoshino's work [8], which gives amongst others a parametrization of the set of all admissible models on any reasonable concrete regularity structure. See [9] for more results on the structure of the space of models and modelled distributions. Note that in the different components u_τ of a modelled distribution also appear in the paracontrolled approach, in which they are involved in the global description of a possible solution, as opposed to their local meaning in the regularity structure setting.

Section 4. This section essentially follows the line of the corresponding results in [53], Section 7 therein. Note that the setting presented here does not allow to take as initial condition a Dirac mass for instance. One needs for that purpose to set a Besov counterpart of the theory, as opposed to the Hölder flavoured version presented here. See Hairer and Labbé's work [58], Hensel and Rosati's work [63], or Singh and Teichmann's work [77].

In a different direction a number of works have been done on quasilinear singular stochastic PDEs [7, 12, 13, 46, 47, 69, 70, 75, 76].

Section 5. The notion of renormalization structures and compatible regularity and renormalization structures introduced in this section is new. It encodes in a simple way the mechanics at work in Bruned, Hairer, and Zambotti's work [20].

Section 6. This section contains the core insights of Bruned, Chandra, Chevyrev, and Hairer's work [17], implemented here on the example of the generalized KPZ equation. The relevance of the notion of pre-Lie algebra was first noticed in the work [18] of Bruned, Chevyrev, and Friz on rough paths. The article [24] would provide a pre-history of the pre-Lie algebra. The comodule-bialgebra structure of the Butcher–Connes–Kreimer Hopf algebra was first investigated in the work [22] of Calaque, Ebrahimi-Fard, and Manchon; it played a key motivating role in the work of Bruned, Chevyrev, and Friz. The description of the free pre-Lie algebra in this setting is due to Chapoton and Livernet [28]. The notion of multi-pre-Lie algebra was introduced in the work [17] of Bruned, Chandra, Chevyrev, and Hairer, where the free multi-pre-Lie was first described.

The $\mathfrak{o}$-decorations introduced here under the form of $\bullet^{n,\alpha}$ are forced by our construction of compatible regularity and renormalization structures for the generalized KPZ equation, given in Section 9. It has no dynamical meaning. Bailleul and Bruned showed in [6] how to obtain the renormalized equation without using extended decorations for a large class of renormalization procedures including the BPHZ scheme. This work is based on the recursive renormalization scheme introduced by Bruned [16].

The setting described here is robust enough to deal with equations driven by multiple noises, or systems of equations driven by multiple noises. We take Funaki's example [45] of the random motion of a rubber on a manifold as an archetype – see also [19, 55]. The unknown u is a spacetime function with values in $\mathbb{R}^d$, solution of the system

$$(\partial_t - \partial_x^2)u = \Gamma(u)\big(\partial_x u, \partial_x u\big) + \Sigma(u)\xi,$$

where $\Gamma(z)$ is a symmetric matrix on $\mathbb{R}^d$, and $\Sigma(z)$ a linear map from $\mathbb{R}^k$ to $\mathbb{R}^d$, for any $z \in \mathbb{R}^d$, and $\xi = (\xi^1, \dots, \xi^k)$ is an k-dimensional tuple of identically distributed independent one-dimensional spacetime white noises. We still have only one operator $(\partial_t - \partial_x^2)$ in this example, so the edge-type set is here the same as in the study of the generalized KPZ equation. The node set is changed from $\{\circ, \bullet\} \times \mathbb{N}^{d+1}$ to $\{\circ^1, \dots, \circ^k, \bullet\} \times \mathbb{N}^{d+1}$ to account for the fact that we have k noises $\xi^1, \dots, \xi^k$ in the system. Things get a bit messier if the system involves different operators, with different regularising properties, and noises with different regularities. The fundamental ideas involved in the analysis remain the same, while the notations needed to take care of this richer setting become heavier. All this is explained in full detail in [20].

Section 7. This section gives what seems to us to be one of the two core results of [20], Theorem 44 here. More general renormalization schemes were introduced by Bruned in [16] and Bailleul and Bruned have shown in [6] how to get back the renormalized equation in a very simple way in a setting with no extended decorations.

Section 8. The fact that the family of solutions to the generalized KPZ equation forms a finite-dimensional manifold of some function space had not been noticed so far. More generally, this is true for any singular stochastic PDE that can be treated by the methods of regularity structures This short section emphasizes that fact.

Section 9. This section builds on the fundamental work [20], with a number of simplifications. The notion of subcritical equation is subtle to check in the general case of a system of equations, as one needs to keep track of how a given symbol of a regularity structure "flows" in the different pieces of a system, involving possibly operators with different regularizing properties. The meaning of subcriticality remains, though.

Funding. The first author was partially supported by the ANR via the ANR-16-CE40-0020-01 Grant. The second author was partially supported by the JSPS KAKENHI Grant No. 19K14556.

References

[1] R. Azencott, Petites perturbations aléatoires des systèmes dynamiques: développements asymptotiques. *Bull. Sci. Math. (2)* **109** (1985), no. 3, 253–308 Zbl 0591.60023 MR 0822827

[2] I. Bailleul, Flows driven by rough paths. *Rev. Mat. Iberoam.* **31** (2015), no. 3, 901–934 Zbl 1352.34089 MR 3420480

[3] I. Bailleul and F. Bernicot, Heat semigroup and singular PDEs. *J. Funct. Anal.* **270** (2016), no. 9, 3344–3452 Zbl 1416.35303 MR 3475460

[4] I. Bailleul and F. Bernicot, High order paracontrolled calculus. *Forum Math. Sigma* **7** (2019), article no. e44 Zbl 1473.60093 MR 4061967

[5] I. Bailleul, F. Bernicot, and D. Frey, Space-time paraproducts for paracontrolled calculus, 3D-PAM and multiplicative Burgers equations. *Ann. Sci. Éc. Norm. Supér. (4)* **51** (2018), no. 6, 1399–1456 Zbl 1430.60053 MR 3940901

[6] I. Bailleul and Y. Bruned, Renormalised singular stochastic PDEs. 2021, arXiv:2101.11949v2

[7] I. Bailleul, A. Debussche, and M. Hofmanová, Quasilinear generalized parabolic Anderson model equation. *Stoch. Partial Differ. Equ. Anal. Comput.* **7** (2019), no. 1, 40–63 Zbl 1448.35320 MR 3916262

[8] I. Bailleul and M. Hoshino, Paracontrolled calculus and regularity structures I. *J. Math. Soc. Japan* **73** (2021), no. 2, 553–595 Zbl 1467.35365 MR 4255075

[9] I. Bailleul and M. Hoshino, Paracontrolled calculus and regularity structures II. *J. Éc. polytech. Math.* **8** (2021), 1275–1328 Zbl 1507.60146 MR 4296495

[10] I. Bailleul and M. Hoshino, Renormalization of random models: A review. 2024, arXiv:2409.15984v1

[11] I. Bailleul and M. Hoshino, Random models on regularity-integrability structures. [v1] 2023, [v2] 2024, arXiv:2310.10202v2

[12] I. Bailleul, M. Hoshino, and S. Kusuoka, Regularity Structures for Quasilinear Singular SPDEs. *Arch. Ration. Mech. Anal.* **248** (2024), no. 6, article no. 127 Zbl 07962825 MR 4833233

[13] I. Bailleul and A. Mouzard, Paracontrolled calculus for quasilinear singular PDEs. *Stoch. Partial Differ. Equ. Anal. Comput.* **11** (2023), no. 2, 599–650 Zbl 1519.60056 MR 4588618

[14] G. Ben Arous, Methods de Laplace et de la phase stationnaire sur l'espace de Wiener. *Stochastics* **25** (1988), no. 3, 125–153 Zbl 0666.60026 MR 0999365

[15] N. Berglund, *An introduction to singular stochastics PDEs—Allen–Cahn equations, metastability, and regularity structures*. EMS Ser. Lect. Math., EMS Press, Berlin, 2022 Zbl 1511.35001 MR 4458524

[16] Y. Bruned, Recursive formulae in regularity structures. *Stoch. Partial Differ. Equ. Anal. Comput.* **6** (2018), no. 4, 525–564 Zbl 1427.35378 MR 3867706

[17] Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer, Renormalising SPDEs in regularity structures. *J. Eur. Math. Soc. (JEMS)* **23** (2021), no. 3, 869–947 Zbl 1465.60057 MR 4210726

[18] Y. Bruned, I. Chevyrev, P. K. Friz, and R. Preiß, A rough path perspective on renormalization. *J. Funct. Anal.* **277** (2019), no. 11, article no. 108283 Zbl 1436.60086 MR 4013830

[19] Y. Bruned, F. Gabriel, M. Hairer, and L. Zambotti, Geometric stochastic heat equations. *J. Amer. Math. Soc.* **35** (2022), no. 1, 1–80 Zbl 1476.60101 MR 4322389

[20] Y. Bruned, M. Hairer, and L. Zambotti, Algebraic renormalisation of regularity structures. *Invent. Math.* **215** (2019), no. 3, 1039–1156 Zbl 1481.16038 MR 3935036

[21] Y. Bruned and D. Manchon, Algebraic deformation for (S)PDEs. *J. Math. Soc. Japan* **75** (2023), no. 2, 485–526 Zbl 1514.35492 MR 4578048

[22] D. Calaque, K. Ebrahimi-Fard, and D. Manchon, Two interacting Hopf algebras of trees: A Hopf-algebraic approach to composition and substitution of B-series. *Adv. in Appl. Math.* **47** (2011), no. 2, 282–308 Zbl 1235.16032 MR 2803804

[23] F. Caravenna and L. Zambotti, Hairer's reconstruction theorem without regularity structures. *EMS Surv. Math. Sci.* **7** (2020), no. 2, 207–251 Zbl 1479.46054 MR 4261666

[24] A. Cayley, On the theory of the analytical forms called trees. *Lond. Edinb. Dubl. Phil. Mag* **13** (1857), no. 85, 172–176

[25] A. Chandra and L. Ferdinand, A flow approach to the generalized KPZ equation. 2024, arXiv:2402.03101v4

[26] A. Chandra and M. Hairer, An analytic BPHZ theorem for regularity structures. [v1] 2016, [v5] 2018, arXiv:1612.08138v5

[27] A. Chandra and H. Weber, Stochastic PDEs, regularity structures, and interacting particle systems. *Ann. Fac. Sci. Toulouse Math. (6)* **26** (2017), no. 4, 847–909 Zbl 1421.81001 MR 3746645

[28] F. Chapoton and M. Livernet, Pre-Lie algebras and the rooted trees operad. *Internat. Math. Res. Notices* (2001), no. 8, 395–408 Zbl 1053.17001 MR 1827084

[29] P. R. Chernoff, Note on product formulas for operator semigroups. *J. Functional Analysis* **2** (1968), 238–242 Zbl 0157.21501 MR 0231238

[30] I. Corwin and H. Shen, Some recent progress in singular stochastic partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* **57** (2020), no. 3, 409–454 Zbl 1448.35003 MR 4108091

[31] G. Da Prato and A. Debussche, Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31** (2003), no. 4, 1900–1916 Zbl 1071.81070 MR 2016604

[32] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*. Encyclopedia Math. Appl. 44, Cambridge University Press, Cambridge, 1992 Zbl 0761.60052 MR 1207136

[33] A. M. Davie, Differential equations driven by rough paths: An approach via discrete approximation. *Appl. Math. Res. Express. AMRX* (2008), article no. ID Zbl 1163.34005 MR 2387018

[34] P. Duch, Flow equation approach to singular stochastic PDEs. [v1] 2021, [v2] 2022, arXiv:2109.11380v2

[35] P. Duch, Renormalization of singular elliptic stochastic PDEs using flow equation. [v1] 2022, [v2] 2023, arXiv:2201.05031v2

[36] P. Duch, M. Gubinelli, and P. Rinaldi, Parabolic stochastic quantisation of the fractional Φ^4_3 model in the full subcritical regime. [v1] 2023, [v3] 2024, arXiv:2303.18112v3

[37] D. Feyel and A. de La Pradelle, Curvilinear integrals along enriched paths. *Electron. J. Probab.* **11** (2006), article no. 34 Zbl 1110.60031 MR 2261056

[38] D. Feyel, A. de La Pradelle, and G. Mokobodzki, A non-commutative sewing lemma. *Electron. Commun. Probab.* **13** (2008), 24–34 Zbl 1186.26009 MR 2372834

[39] H. Figueroa and J. M. Gracia-Bondía, The uses of Connes and Kreimer's algebraic formulation of renormalization theory. *Internat. J. Modern Phys. A* **19** (2004), no. 16, 2739–2754 Zbl 1080.81046 MR 2062903

[40] L. Foissy, Commutative and non-commutative bialgebras of quasi-posets and applications to Ehrhart polynomials. *Adv. Pure Appl. Math.* **10** (2019), no. 1, 27–63 Zbl 1407.16032 MR 3895621

[41] L. Foissy, Algebraic structures on typed decorated rooted trees. *SIGMA Symmetry Integrability Geom. Methods Appl.* **17** (2021), article no. 086 Zbl 1504.17002 MR 4315488

[42] P. K. Friz, P. Gassiat, and P. Pigato, Precise asymptotics: Robust stochastic volatility models. *Ann. Appl. Probab.* **31** (2021), no. 2, 896–940 Zbl 1476.60183 MR 4254499

[43] P. K. Friz and M. Hairer, *A course on rough paths. With an introduction to regularity structures*. Universitext, Springer, Cham, 2020 Zbl 1437.60002 MR 4174393

[44] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths. Theory and applications*. Cambridge Stud. Adv. Math. 120, Cambridge University Press, Cambridge, 2010 Zbl 1193.60053 MR 2604669

[45] T. Funaki, A stochastic partial differential equation with values in a manifold. *J. Funct. Anal.* **109** (1992), no. 2, 257–288 Zbl 0768.60055 MR 1186323

[46] M. Furlan and M. Gubinelli, Paracontrolled quasilinear SPDEs. *Ann. Probab.* **47** (2019), no. 2, 1096–1135 Zbl 1447.60099 MR 3916943

[47] M. Gerencsér and M. Hairer, A solution theory for quasilinear singular SPDEs. *Comm. Pure Appl. Math.* **72** (2019), no. 9, 1983–2005 Zbl 1458.60077 MR 3987723

[48] M. Gubinelli, Controlling rough paths. *J. Funct. Anal.* **216** (2004), no. 1, 86–140 Zbl 1058.60037 MR 2091358

[49] M. Gubinelli, Ramification of rough paths. *J. Differential Equations* **248** (2010), no. 4, 693–721 Zbl 1315.60065 MR 2578445

[50] M. Gubinelli, P. Imkeller, and N. Perkowski, Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3** (2015), article no. e6 Zbl 1333.60149 MR 3406823

[51] M. Gubinelli and N. Perkowski, *Lectures on singular stochastic PDEs*. Ensaios Mat. 29, Sociedade Brasileira de Matemática, Rio de Janeiro, 2015 Zbl 1337.60005 MR 3445609

[52] M. Gubinelli and N. Perkowski, An introduction to singular SPDEs. In *Stochastic partial differential equations and related fields*, pp. 69–99, Springer Proc. Math. Stat. 229, Springer, Cham, 2018 Zbl 1405.60092 MR 3828162

[53] M. Hairer, A theory of regularity structures. *Invent. Math.* **198** (2014), no. 2, 269–504 Zbl 1332.60093 MR 3274562

[54] M. Hairer, Introduction to regularity structures. *Braz. J. Probab. Stat.* **29** (2015), no. 2, 175–210 Zbl 1316.81061 MR 3336866

[55] M. Hairer, The motion of a random string. 2016, arXiv:1605.02192v2

[56] M. Hairer, Renormalisation of parabolic stochastic PDEs. *Jpn. J. Math.* **13** (2018), no. 2, 187–233 Zbl 1430.60055 MR 3855740

[57] M. Hairer, Solving the KPZ equation. *Ann. of Math. (2)* **178** (2013), no. 2, 559–664 Zbl 1281.60060 MR 3071506

[58] M. Hairer and C. Labbé, The reconstruction theorem in Besov spaces. *J. Funct. Anal.* **273** (2017), no. 8, 2578–2618 Zbl 1380.46031 MR 3684891

[59] M. Hairer and J. Mattingly, The strong Feller property for singular stochastic PDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2018), no. 3, 1314–1340 Zbl 1426.60088 MR 3825883

[60] M. Hairer and P. Schönbauer, The support of singular stochastic partial differential equations. *Forum Math. Pi* **10** (2022), article no. e1 Zbl 1490.60186 MR 4370580

[61] M. Hairer and R. Steele, The BPHZ theorem for regularity structures via the spectral gap inequality. *Arch. Ration. Mech. Anal.* **248** (2024), no. 1, article no. 9 Zbl 1546.60198 MR 4693281

[62] M. Hairer and H. Weber, Large deviations for white-noise driven, nonlinear stochastic PDEs in two and three dimensions. *Ann. Fac. Sci. Toulouse Math. (6)* **24** (2015), no. 1, 55–92 Zbl 1319.60141 MR 3325951

[63] S. Hensel and T. Rosati, Modelled distributions of Triebel–Lizorkin type. *Studia Math.* **252** (2020), no. 3, 251–297 Zbl 1460.46024 MR 4069993

[64] M. Hoshino, A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate. 2023, arXiv:2310.07396v1

[65] M. Hoshino and R. Takano, A semigroup approach to the reconstruction theorem and the multilevel Schauder estimate for singular modelled distributions. 2024, arXiv:2408.04322v2

[66] Y. Inahama and H. Kawabi, Asymptotic expansions for the Laplace approximations for Itô functionals of Brownian rough paths. *J. Funct. Anal.* **243** (2007), no. 1, 270–322 Zbl 1114.60062 MR 2291439
[67] A. Kupiainen, Renormalization group and stochastic PDEs. *Ann. Henri Poincaré* **17** (2016), no. 3, 497–535 Zbl 1347.81063 MR 3459120
[68] A. Kupiainen and M. Marcozzi, Renormalization of generalized KPZ equation. *J. Stat. Phys.* **166** (2017), no. 3-4, 876–902 Zbl 1369.82011 MR 3607594
[69] P. Linares, F. Otto, and M. Tempelmayr, The structure group for quasi-linear equations via universal enveloping algebras. *Commun. Am. Math. Soc.* **3** (2023), 1–64 Zbl 1535.60156 MR 4537770
[70] P. Linares, F. Otto, M. Tempelmayr, and P. Tsatsoulis, A diagram-free approach to the stochastic estimates in regularity structures. *Invent. Math.* **237** (2024), no. 3, 1469–1565 Zbl 07895067 MR 4777092
[71] C. Liu, D. J. Prömel, and J. Teichmann, Stochastic analysis with modelled distributions. *Stoch. Partial Differ. Equ. Anal. Comput.* **9** (2021), no. 2, 343–379 Zbl 1470.35448 MR 4261328
[72] T. J. Lyons, Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** (1998), no. 2, 215–310 Zbl 0923.34056 MR 1654527
[73] D. Manchon, Hopf algebras, from basics to applications to renormalization. [v1] 2004, [v2] 2006, arXiv:math/0408405v2
[74] J. Martin, *Refinements of the solution theory for singular SPDEs*. Ph.D. thesis, 2018
[75] F. Otto, J. Sauer, S. Smith, and H. Weber, Parabolic equations with rough coefficients and singular forcing. 2018, arXiv:1803.07884v2
[76] F. Otto and H. Weber, Quasilinear SPDEs via rough paths. *Arch. Ration. Mech. Anal.* **232** (2019), no. 2, 873–950 Zbl 1426.60090 MR 3925533
[77] H. Singh and J. Teichmann, An elementary proof of the reconstruction theorem. 2018, arXiv:1812.03082v1
[78] M. E. Sweedler, *Hopf algebras*. Math. Lect. Note Ser., W. A. Benjamin, New York, 1969 Zbl 0194.32901 MR 0252485

Received 29 March 2023; revised 26 October 2024.

Ismaël Bailleul
Laboratoire de Mathématiques de Bretagne Atlantique, Université de Bretagne Occidentale, 6 avenue le Gorgeu, LMBA-UMR 6205, 29238 Brest, France; ismael.bailleul@univ-brest.fr

Masato Hoshino
Graduate School of Engineering Science, Osaka University, Osaka 560-8531, Japan; hoshino@sigmath.es.osaka-u.ac.jp