arXiv:2510.20716v1 [math.AP] 23 Oct 2025

Large field problem in coercive singular PDEs

Ilya Chevyrev* Massimiliano Gubinelli†

October 24, 2025

Abstract

We derive a priori estimates for singular differential equations of the form

$$\mathcal{L}\varphi = P(\varphi, \nabla\varphi) + f(\varphi, \nabla\varphi)\xi$$

where P is a polynomial, f is a sufficiently well-behaved function, and ξ is an irregular distribution such that the equation is subcritical. The differential operator $\mathcal{L}$ is either a derivative in time, in which case we interpret the equation using rough path theory, or a heat operator, in which case we interpret the equation using regularity structures. Our only assumption on P is that solutions with $\xi = 0$ exhibit coercivity. Our estimates are local in space and time, and independent of boundary conditions.

One of our main results is an abstract estimate that allows one to pass from a local coercivity property to a global one using scaling, for a large class of equations. This allows us to reduce the problem of deriving a priori estimates to the case when ξ is small.

Contents

1 Introduction **2**
1.1 Notation 6

2 Main idea in classical setting **9**

3 Abstract bounds **13**

4 Applications to well-posed equations **16**
4.1 Classical case revisited 16
4.2 Young ODEs 17

5 Rough paths **22**
5.1 General a priori estimates 22
5.2 Simple linear condition on forcing 24
5.3 Proof of Theorem 5.5 25

6 Singular PDEs **29**
6.1 Setup and regularity structures 30
6.2 Main result 40
6.3 Examples: linear conditions on f 41
6.4 Setup for proof 45

*SISSA, Trieste, 34136, Italy. ichevyrev@gmail.com
†Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK. gubinelli@maths.ox.ac.uk

6.5 Scaling 46
6.6 Structure of elementary differentials 49
6.7 Local reconstruction bounds 54
6.8 Integration bounds 55
6.9 Remainder bounds 58
6.10 Local coercivity 62

A Extension of distributions **64**

1 Introduction

Global aspects of solution theory to stochastic partial differential equations (SPDEs) remain an open problem. In particular, the large field regime—where solutions may become arbitrarily large—poses significant challenges for a priori estimates and existence.

We develop a strategy for a priori estimates based on the fact that the obstruction to such estimates is given by the presence of solutions which blow up at specific space-time points. In the neighbourhood of such points, the solution is large and the equation can be approximated by its scale-invariant version. In the subcritical regime, this corresponds to a semi-classical situation where the noise become small and the equation is dominated by its deterministic part. If the deterministic equation has a coercive non-linearity then blow-up cannot occur. This intuition can be made rigorous and we show that it leads to a priori estimates for a large class of subcritical singular SPDEs.

We note an intriguing convergence of ideas, where one begins with an example that violates a desired estimate (or property) and, through a scaling argument, arrives at a contradiction by reducing the problem to a much simpler case—typically in a 'semi-classical' regime where the noise is negligible.

In particular, we are aware of two other situations where this happens:

- **Symmetry:** In [CCHS22, CCHS24, BGHZ22] the authors use this line of reasoning to prove structural properties of renormalisation, in particular the equivariance under symmetries of the renormalized equation. It were these works that inspired us to tackle the large field problem via a similar strategy.

- **Regularity:** This line of reasoning is used to prove Schauder estimates in a kernel-free way, see e.g. [Sim97, SS25, EW24]. The idea is to zoom into a space-time point where the desired estimate fails and then use scaling to reduce the problem to a situation where the contributions by the noise are small. In this regime, the desired estimate can be proved by a Liouville property of the linear part of the equation.

In all these cases, we see that a particular property of the SPDE is related to some limit where the noise disappears, either in small scales or for large values of the field, i.e. near an eventual blow-up, or in a regime where a symmetry is broken at the semi-classical level.

It would be interesting to explore this heuristic further especially in connection with geometric aspects of singular SPDEs.

Setting and main results. We consider a partial differential equation (PDE) of the form

$$\mathcal{L}\varphi = P(\varphi, \nabla\varphi) + f(\varphi, \nabla\varphi)\xi \,. \tag{1.1}$$

Here, $\mathcal{L}$ is the heat operator $\partial_t - \Delta = \partial_1 - \sum_{i=2}^{d} \partial_i^2$ on $\mathbb{R}^d$, P is a polynomial of φ and its spatial derivative $\nabla\varphi = (\partial_i\varphi)_{i=2}^d$, and f is a function of $(\varphi, \nabla\varphi)$. We allow the case $d = 1$, in which case $\mathcal{L} = \partial_t$, so the equation is an ordinary differential equation (ODE) and P and f depend

only on φ. The equation is posed for a (vector-valued) function $\varphi\colon \Omega \to \mathbb{R}^m$, where $m \geq 1$ and $\Omega = (-1,0) \times (-1,1)^{d-1}$ is the backwards parabolic ball.

Here ξ is a distribution in a Hölder–Besov space $\mathcal{C}^\beta(\Omega, \mathbb{R}^n)$ with $\beta \leq 0$ (like φ, we allow ξ to be vector-valued). We are especially interested in the case where (1.1) is *singular*, meaning that β is sufficiently negative to render the equation classically ill-posed. In this case, we interpret (1.1) either using rough paths [Lyo98] (if $d = 1$) or using regularity structures [Hai14] (if $d \geq 2$). In both cases, to give meaning to the solution, we require a rough path / model Z above ξ.

Our assumptions regarding the equation can be stated roughly as follows:

- (1.1) is scaling subcritical in the sense of [Hai14].
- The equation $\mathcal{L}\varphi = P(\varphi, \nabla\varphi)$ is scale-invariant with exponent $\alpha > 0$. Moreover, there exists $\delta > 0$ such that, if $\mathcal{L}\varphi = P(\varphi, \nabla\varphi)$ on Ω with $\|\varphi\|_\infty \leq 1$, then $|\varphi(0)| < 1 - \delta$.
- $f\colon \mathbb{R}^m \to L(\mathbb{R}^n, \mathbb{R}^m)$ has sufficient regularity and growth at infinity.

The first condition is a bound on the regularity of ξ and is necessary to apply the theories of rough paths or regularity structures. For example, for $d \geq 2$, if $P = -|\varphi|^2\varphi$ and f is constant, then we require $\beta > -3$, and if f is not constant but does not depend on $\nabla\varphi$, then we require $\beta > -2$. The condition roughly states that the solution φ is, on small scales, a perturbation of the solution to the linear equation $\mathcal{L}\psi = \xi$ (in the language of QFT, it corresponds to super-renormalisability).

In the second condition, 'scale-invariance' means that, if φ solves $\mathcal{L}\varphi = P(\varphi, \nabla\varphi)$, then

$$\psi(t,x) = \lambda^\alpha \varphi(z + (\lambda^2 t, \lambda x)) \tag{1.2}$$

solves the same equation for all $\lambda > 0$ and $z \in \mathbb{R}^d$. For example, we can take P of the form $P(\varphi, \nabla\varphi) = \varphi^p + \varphi^{\frac{p-1}{2}}\nabla\varphi$ and the corresponding exponent is $\alpha = \frac{2}{p-1}$. (The condition implies that P depends on $\nabla\varphi$ in an affine way.) By a scaling argument, the second part of the condition is actually equivalent to the much stronger 'coming down from infinity' property which states that $|\varphi(0)| \lesssim 1$ whenever φ solves $\mathcal{L}\varphi = P(\varphi, \nabla\varphi)$ on Ω (uniformly in the boundary conditions). The fact that P is a polynomial is not crucial for our argument and we make this assumption only to simplify the proof.

We leave the conditions on f intentionally vague for now, as their formulation is lengthy and depends on whether we are in the ODE case ($d = 1$) or the PDE case ($d \geq 2$). We highlight, however, that our regularity assumptions do not impose uniqueness of solutions to (1.1).

Under these conditions, we derive a priori estimates for any solution φ of the remainder equation of (1.1). By 'solution of the remainder equation' we mean that φ is the function-like remainder of the 'full' solution. In the case of rough paths or when f is not constant, φ is already the full solution. But if f is constant and $d \geq 2$, then, like in e.g. [BCCH21], we write the 'full' solution as $\bar{\varphi} = X + \varphi$, where X is an explicit polynomial function of ξ that, in general, is a distribution not in L^∞. Since we derive bounds in L^∞, we focus henceforth on φ.

One heuristic formulation of our main results is

Theorem 1.1. *Suppose $\varphi\colon \Omega \to \mathbb{R}^m$ solves the remainder equation of (1.1) with input rough path / model Z above ξ. Then*

$$|\varphi(z)| \lesssim \max\{[\![f, Z]\!]_{z,\lambda}, |\!|\!|z|\!|\!|^{-\alpha}\} \tag{1.3}$$

where $|\!|\!|z|\!|\!|$ is the distance of z to the parabolic boundary of Ω and $[\![f, Z]\!]_{z,\lambda}$ is a 'norm' of the pair (f, Z) that depends on the growth of the derivatives of f at infinity and on the size of Z in a ball centred at z of radius $\lambda \leq \frac{1}{2}|\!|\!|z|\!|\!|$.

We give a precise and sharper formulation for the rough path case in Theorem 5.5 and for the regularity structure case in Theorem 6.35. The conditions on f and P are made precise in the respective sections.

Method of proof. Our proof of Theorem 1.1 is based on a scaling argument and proceeds in two steps.

First, we suppose the pair (f, Z) is suitably 'small' and that $\|\varphi\|_{\infty;\Omega} \leq 1$. Recalling $\delta > 0$ from our assumption on P, we prove in this case that $|\varphi(0)| < 1 - \delta/2$. This step is significantly easier than handling (1.1) right away because smallness of $\|\varphi\|_\infty$ and (f, Z) can be used to show $\|\varphi - \mathring{\varphi}\|_\infty < \delta/2$, where $\mathring{\varphi}$ solves $\mathcal{L}\mathring{\varphi} = P(\mathring{\varphi}, \nabla\mathring{\varphi})$ with the same boundary conditions as φ. Then $|\mathring{\varphi}| < 1 - \delta$ by assumption, which leads to the claimed bound $|\varphi(0)| < 1 - \delta/2$. At this point, we have not used scale-invariance of the equation.

Second, no longer assuming any smallness, we rescale $\varphi \mapsto \psi$ according to the transformation (1.2) around a point z that maximises the ratio of the left- to right-hand side of (1.3). We choose $\lambda < 1$ in (1.2) such that $|\psi(0)|_{\infty;\Omega} = 1 - \delta/2$ and remark that ψ solves the same equation except (f, Z) becomes rescaled. Due to subcriticality and our growth assumptions on f, this rescaled version of (f, Z) becomes 'small'. This implies $\|\psi\|_{\infty;\Omega} > 1$ by the first step, which readily leads to a contradiction on the maximality of z.

To illustrate the idea in a simple setting, we implement this strategy in Section 2 for the PDE $\mathcal{L}\varphi = -\varphi^p + \xi$ in the classically well-posed regime $\beta > -2$.

The argument in the second step is rather general and does not rely on the fact that we work with differential equations. We therefore find it useful to state it as an abstract local-to-global coercivity bound, a simplified and heuristic version of which is as follows.

Suppose we are given

- a metric space (Σ, ϱ) with a 'boundary' $\partial\Sigma \subset \Sigma$ and a distinguished point $o \in \Sigma$,
- an embedding $T_{z,\lambda} \colon \Sigma \to \Sigma$ for every $z \in \Sigma$ and $0 < \lambda \leq |\!|\!| z |\!|\!|$, where $|\!|\!| z |\!|\!| = \inf_{y \in \partial\Sigma} \varrho(y, z)$, such that $T_{z,\lambda}(o) = z$,
- a set of 'drivers' $\mathbb{M}$ with a 'norm' $[\![\,\cdot\,]\!] \colon \mathbb{M} \to [0, \infty)$ and 'scaling map' $R_{z,\lambda} \colon \mathbb{M} \to \mathbb{M}$ such that $[\![R_{z,\lambda} M]\!] \leq \lambda^\alpha [\![M]\!]$ for some $\alpha > 0$, and
- a set of 'solutions' $\mathbb{S}_M$ for every $M \in \mathbb{M}$, elements of which are functions $\varphi \colon \Sigma \to \mathbb{R}^m$, such that $\lambda^\alpha \varphi \circ T_{z,\lambda} \in \mathbb{S}_{R_{z,\lambda} M}$ for all $\varphi \in \mathbb{S}_M$.

Finally, suppose there exists $\delta > 0$ such that, if $[\![M]\!] \leq 1$ and $\|\varphi\|_\infty \leq 1 + \delta$, then

$$|\varphi(o)| < 1 \,.$$

We interpret this final assumption as a *local coercivity* condition.

Theorem 1.2. *In the above setting, every $\varphi \in \mathbb{S}_M$ satisfies*

$$|\varphi(z)| \lesssim \max\{[\![M]\!], |\!|\!| z |\!|\!|^{-\alpha}\} \,.$$

We refer to Section 3, especially Theorem 3.4, for a precise formulation, and to Section 4 for simple applications of this result to Young ODEs.

We believe this scaling argument is quite general and can be adapted to other settings, such as to discrete systems. One can also adapt it to incorporate boundary data in the estimates, but we refrain from doing so in this work as it complicates the abstract setup.

Several further remarks are in order.

- **Relation to SPDEs.** Our results are inspired by recent developments in *stochastic* PDEs. Indeed, many interesting examples of (1.1) arise from taking ξ as a *random* distribution (e.g. white noise). In applications of rough paths / regularity structures to such equations, one would construct a random rough path / model above ξ. See [CH16, LOTT24, HS24] for systematic ways to build renormalised random models above a class of distributions via smooth approximations, and [FV10, FH20] for the case of rough paths. In the case of regularity structures, for these smooth, renormalised models, the corresponding solution to (1.1) solves a *renormalised* PDE, see [BCCH21]. The results of the present article are entirely deterministic and we only require a rough path / model as input.

- **Notion of solution.** In the case of regularity structures, our notion of solution to (1.1) is somewhat new (see Definition 6.32) and is inspired by Davie's notion of solution to rough differential equations [Dav08], which we recall in Section 5. Our notion of solution is consistent with that of [Hai14], but it avoids making sense of calculus operations on the space of modelled distributions. We also do not require φ to solve a (possibly renormalised) PDE. (We note that a similar notion was independently proposed by [BOS25] for the Φ^4 equation in the multi-index approach to regularity structures.)

- **Local theory.** Our abstract formulation of Theorem 1.2 is indifferent to the underlying solution theory of the PDE. The only part where we use regularity structures is in the verification of the local coercivity condition, and we believe that, for certain equations, one could verify the condition with other methods, such as paracontrolled analysis [GIP15], renormalisation group and flow approach [Kup16, Duc25], or the the multi-index approach to regularity structures [OW19, LOT23, OSSW25]. We believe it would be of interest if our scaling argument can be combined with other local solutions theories to broaden the scope of equations covered (e.g. to quasi-linear equations).

- **Refinements of regularity structures.** As part of our argument, we make small refinements of standard estimates in regularity structures. This includes the reconstruction theorem (Lemma 6.61) and multi-scale Schauder estimates (Lemma 6.63), for which we prove bounds that are uniform in the size of the truncation of the integration kernel. We also point out that, to bypass issues related to non-uniqueness of reconstructions of singular modelled distributions, we show that a distribution $\xi \in \mathcal{D}'(B)$ which is in $\mathcal{C}^\eta$ in an open ball $B \subset \mathbb{R}^d$ admits an extension to $\mathcal{C}^\eta(\mathbb{R}^d)$ with support contained in the closure of B, see Appendix A – this result is similar to several other extension theorems in the literature, but our proof appears different and rather elementary.

Comparison to other results. Space-time local bounds of the form (1.3) have previously been considered for special cases of (1.1) in [MW20a, MW20b, CMW23, BCMW22, JP25], and our formulation of Theorem 1.1 is very much inspired by these works.

Our estimates are more general in the sense that they contain those in the aforementioned works as special cases. Specifically, [MW20a] considers the PDE $\mathcal{L}\varphi = -\varphi^p + g(\varphi) + \xi$, where g is bounded and $\xi \in \mathcal{C}^\beta$ with $\beta > -2$, i.e. the classically well-posed regime, and we essentially recover this case in Section 2. In [MW20b, CMW23], the authors study the PDE $\mathcal{L}\varphi = -\varphi^p + \xi$ in the full subcritical regime $\beta > -3$ and [JP25] analyses $\mathcal{L}\varphi = -\varphi^2 + \varphi\xi$ in the first singular regime $\xi \in \mathcal{C}^{-1-\kappa}$ for $\kappa > 0$ small. Section 6 generalises these results. Finally, [BCMW22] considers the rough path case ($d = 1$) where f either behaves like a polynomial or has bounded derivatives of all orders. Our estimates in Section 5 agree in these two cases but we treat more general forms of f (e.g. we prove stronger bounds for f of the form $f(\varphi) = \varphi^q \sin(\varphi^p)$).

We believe that the major advantage of our approach is its simplicity and its applicability. For example, we avoid delicate applications of the maximum principle, so are able to treat

vector-valued and scalar-valued equations on an equal footing ([MW20a, MW20b, CMW23] work only in the scalar-valued case, though one does strongly expect that their method can be extended to the vector-valued case with extra efforts). Additionally, we are able to work with the original formulation of models and modelled distributions due to Hairer [Hai14], and only require minor refinements of the reconstruction and integration bounds. Relatedly, we require no qualitative smoothness assumption on ξ (or its lift to a model) and do not need to consider renormalised PDEs. Furthermore, our method applies to stochastic quantisation equations of tilted measures as considered in [HS22] and we believe can be used to show optimal integrability of the invariant measure of $\mathcal{L}\varphi = -\varphi^p + \xi$ in the full subcritical regime.

Open problems.

- **Optimality.** We suspect that for singular ODEs (rough paths), our results are sharp in terms of the scaling exponents of f and ξ. Furthermore, for singular PDEs (regularity structures), we suspect our results are also sharp for *polynomial* f. However, there is a gap between our assumptions for ODEs and PDEs in the case of *non-polynomial* f. E.g. we can leverage decay of f at infinity to get improved bounds for ODEs, but do not know if this is possible for PDEs. Additionally, if f and its derivatives are bounded, then, in the case of PDEs, we require P to have sufficiently large degree depending on the regularity of ξ, while there is no restriction on the degree of P in the case of ODEs; compare Assumption 6.22 with Assumptions 5.3-5.4 and Section 5.2, see also Remark 6.23. We know of cases where the exponents in our bounds for PDEs are not sharp but did not find a way to systematically improve the bounds in general. We find it an interesting problem to determine if the gaps between the two cases can be closed.

- **Role of locality.** In the way we apply our general scaling principle, we rely on the locality of the heat operator $\mathcal{L}$. It would be interesting to find extensions of the method that are able to handle non-local equations, such as the stochastic quantisation equation of the fractional Φ^4 model analysed recently in [EW24, DGR23].

- **Role of coercivity.** Our argument works for equations with super-linear damping terms (such as Φ^p models), for which we show the space-time localisation bound (1.3). There are singular stochastic PDEs of interest for which global well-posedness has been shown but no space-time localisation bound is expected, see for example [BC24b, CS23] for the stochastic 2D Abelian Higgs and Yang–Mills equations, [BC24a, CdW24, SZZ25] for the sine-Gordon and gPAM models, and [HR24, HZ25] for the stochastic Navier–Stokes equations. There are also equations of interest for which only local but no global solution theory is established, such as the 3D stochastic Yang-Mills equations [CCHS24] and the geometric KPZ equation [BGHZ22]. We find it an important problem to understand if scaling methods can be applied to such equations that lack super-linear damping terms.

1.1 Notation

We collect some commonly used notation. Denote $\mathbb{N} = \{0, 1, \dots\}$. For $k \in \mathbb{N}$, denote $[k] = \{1, \dots, k\}$, understood as the empty set if $k = 0$.

For a set S, we interchangeably treat elements $\mathbf{k} \in \mathbb{N}^S$ both as a function $\mathbf{k}\colon S \to \mathbb{N}$ and as a multi-set $\mathbf{k} = (k_1, \dots, k_{|\mathbf{k}|})$ with elements $k_i \in S$, where we denote by $|\mathbf{k}|$ the total number of elements in $\mathbf{k}$. We denote $\mathbf{k}! = \prod_{k \in S} \mathbf{k}(k)!$.

We write $\mathbf{1}_A$ for the indicator function of a set A. For a function $\xi\colon A \to E$, where E is a normed space, we denote

$$\|\xi\|_{\infty;A} = \sup_{a \in A} |\xi(a)| \,.$$

We will sometimes use the shorthand $\xi_a = \xi(a)$ for $a \in A$.

Let $\lfloor \alpha \rfloor \in \mathbb{Z}$ and $\lceil \alpha \rceil \in \mathbb{Z}$ denote the floor and roof of $\alpha \in \mathbb{R}$ respectively. Unless otherwise specified, for $\alpha > 0$, we let $\mathcal{C}^\alpha$ denote the space of N-times differentiable functions, where $N = \lceil \alpha \rceil - 1 \in \mathbb{N}$, whose N-th derivative is γ-Hölder continuous, where $\gamma = \alpha - N \in (0, 1]$. We let $\mathcal{C}^\infty$, $\mathcal{C}_c^\infty$, and $\mathcal{C}$ denote the space of smooth, smooth and compactly supported, and continuous functions respectively. Unless otherwise specified or clear from the context, the target space of functions is $\mathbb{R}$.

We write $f * g$ for the convolution $f * g(x) = \int f(x-y)g(y)\,\mathrm{d}y$ of two functions f, g defined on suitable subsets of $\mathbb{R}^d$.

We write $X \lesssim Y$ and $X \ll Y$ to indicate $X \le CY$ and $CX \le Y$ respectively for some $C \ge 1$ sufficiently large whose dependence on different parameters is made explicit or is clear from the context.

For vector spaces X, Y, we write $L(X, Y)$ for the space of linear maps $X \to Y$.

1.1.1 Scaling

Let $d \ge 1$ and consider scaling $\mathfrak{s} = (\mathfrak{s}_1, \dots, \mathfrak{s}_d) \in \mathbb{N}^d$ with $\mathfrak{s}_i \in [1, \infty)$ for all $i \in [d]$. Denote $|\mathfrak{s}| = \mathfrak{s}_1 + \cdots + \mathfrak{s}_d$ and, for $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, define

$$|x|_{\mathfrak{s}} = \max_{i \in [d]} |x_i|^{1/\mathfrak{s}_i} .$$

We equip $\mathbb{R}^d$ with the corresponding metric $(x, y) \mapsto |x - y|_{\mathfrak{s}}$. We also equip $\mathbb{R}^d$ with multiplication by $\lambda > 0$ defined by

$$\lambda \cdot_{\mathfrak{s}} x = (\lambda^{\mathfrak{s}_1} x_1, \dots, \lambda^{\mathfrak{s}_d} x_d)$$

for which $|\lambda \cdot_{\mathfrak{s}} x|_{\mathfrak{s}} = \lambda |x|_{\mathfrak{s}}$.

For $z \in \mathbb{R}^d$ and $\lambda > 0$, define the affine bijection $T_{z,\lambda} \colon \mathbb{R}^d \to \mathbb{R}^d$ by

$$T_{z,\lambda}(y) = z + \lambda \cdot_{\mathfrak{s}} y , \tag{1.4}$$

and for $\psi \colon \mathbb{R}^d \to \mathbb{R}$ denote

$$\psi_z^\lambda(y) = \lambda^{-|\mathfrak{s}|} \psi(\lambda^{-1} \cdot_{\mathfrak{s}} (y - z)) = \lambda^{-|\mathfrak{s}|} \psi(T_{z,\lambda}^{-1}(y)) .$$

Define the open unit ball

$$\mathfrak{B}(x, h) = \{ y \in \mathbb{R}^d \,:\, |y - x|_{\mathfrak{s}} < h \} . \tag{1.5}$$

For $r \in \mathbb{N}$, denote

$$\mathcal{B}^r = \{ \psi \in \mathcal{C}_c^\infty(\mathfrak{B}(0, 1)) \,:\, \|D^k \psi\|_\infty \le 1 \text{ for all } |k|_{\mathfrak{s}} \le r \} . \tag{1.6}$$

We sometimes call ψ_x^λ, where $\psi \in \mathcal{B}^r$ and $x \in \mathbb{R}^d$, a *test function at scale* $\lambda > 0$.

1.1.2 Parabolic case

In Sections 2 and 6, where we work with parabolic PDEs, we use the following notation. Fix $d \ge 2$ an integer and let $\mathcal{L} = \partial_t - \Delta = \partial_1 - \sum_{i=2}^d \partial_i^2$ be the heat operator on $\mathbb{R}^d$. We set

$$\mathfrak{s} = (\mathfrak{s}_1, \dots, \mathfrak{s}_d) = (2, 1, \dots, 1) . \tag{1.7}$$

Define the backwards parabolic open ball and its parabolic closure

$$\Omega = (-1, 0) \times (-1, 1)^{d-1} \subset \mathbb{R}^d , \qquad \bar{\Omega} = [-1, 0) \times [-1, 1]^{d-1} .$$

We denote the corresponding parabolic boundary by

$$\partial\Omega = \{z \in \bar{\Omega} \,:\, |z|_{\mathfrak{s}} = 1\} \,.$$

We also write $\operatorname{cl}\mathfrak{K}$ for the topological closure of a subset $\mathfrak{K} \subset \mathbb{R}^d$, so that

$$\operatorname{cl}\Omega = [-1,0] \times [-1,1]^{d-1} \,.$$

(Note that $\bar{\Omega} \neq \operatorname{cl}\Omega$ and similarly $\partial\Omega$ differs from the topological boundary of $\Omega \subset \mathbb{R}^d$.)

Define also the rescaled and shifted open parabolic ball

$$B_z(\lambda) = T_{z,\lambda}(\Omega) \,.$$

Distributions. For an open set $\mathfrak{K} \subset \mathbb{R}^d$, let $\mathcal{D}'(\mathfrak{K})$ denote the space of distributions on $\mathfrak{K}$. For $\xi \in \mathcal{D}'(\mathfrak{K})$, $\beta \in \mathbb{R}$, $h > 0$, $r \in \mathbb{N}$, and $x \in \mathfrak{K}$ such that $B_x(h) \subset \mathfrak{K}$, define

$$\|\xi\|_{\beta;x;h;r} = \sup_{\psi\in\mathcal{P}^r} \sup_{\lambda\in(0,h)} |\lambda^{-\beta}\langle \xi, \psi_x^\lambda\rangle| \,, \tag{1.8}$$

where

$$\mathcal{P}^r = \{\psi \in \mathcal{C}_c^\infty(\Omega) \,:\, \|D^k\psi\|_\infty \leq 1 \text{ for all } |k|_{\mathfrak{s}} \leq r\} \subset \mathcal{B}^r \,.$$

For $\psi \in \mathcal{P}^r$, we sometimes call ψ_x^λ a *non-anticipative* test function at scale $\lambda > 0$. If $r = \max\{0, \lfloor 1-\beta \rfloor\}$ and $h = 1$, we drop r, h from the notation and write

$$\|\xi\|_{\beta;x;h} = \|\xi\|_{\beta;x;h;r} \,, \qquad \|\xi\|_{\beta;x} = \|\xi\|_{\beta;x;1} \,.$$

For $\beta \leq 0$, let $\mathcal{C}^\beta(\mathfrak{K})$ denote the space of distributions $\xi \in \mathcal{D}'(\mathfrak{K}^1)$, where

$$\mathfrak{K}^1 = \cup_{x\in\mathfrak{K}} B_x(1) \tag{1.9}$$

(the 1-fattening of $\mathfrak{K}$ using backwards parabolic balls), equipped with the semi-norm

$$\|\xi\|_{\mathcal{C}^\beta(\mathfrak{K})} \stackrel{\text{def}}{=} \sup_{x\in\mathfrak{K}} \|\xi\|_{\beta;x} \,.$$

For $\beta \leq 0$, $\xi \in \mathcal{D}'(\mathfrak{K})$, $x \in \mathbb{R}^d$, and $h > 0$ such that $T_{x,h}(\Omega^1) \subset \mathfrak{K}$, define further

$$\|\xi\|_{\mathcal{C}^\beta;x;h} = \sup_{y\in B_x(h)} \|\xi\|_{\beta;y;h} \,, \qquad \|\xi\|_{\mathcal{C}^\beta;x} = \|\xi\|_{\mathcal{C}^\beta;x;1} \,.$$

Note that $B_y(h) \subset T_{x,h}(\Omega^1) \subset \mathfrak{K}$ for all $y \in B_x(h)$, so that $\|\xi\|_{\beta;y;h}$ indeed makes sense.

Observe that, for all for $\lambda, \delta > 0$,

$$\langle \xi \circ T_{z,\lambda}, \psi_x^\delta\rangle = \langle \xi, \psi^{\lambda\delta}_{T_{z,\lambda}(x)}\rangle$$

and therefore, for $\beta \in \mathbb{R}$,

$$\|\xi \circ T_{z,\lambda}\|_{\beta;x;h;r} = \lambda^\beta \|\xi\|_{\beta;T_{z,\lambda}(x);\lambda h;r} \tag{1.10}$$

and, for $\beta \leq 0$,

$$\|\xi \circ T_{z,\lambda}\|_{\mathcal{C}^\beta(\Omega)} = \lambda^\beta \|\xi\|_{\mathcal{C}^\beta;z;\lambda} \,. \tag{1.11}$$

2 Main idea in classical setting

In this section, we demonstrate the main idea of our approach in the simple setting of classically well-posed non-linear heat equations. We use notation as in Section 1.1.2.

Let $p \in (1,\infty)$ and, for $x \in \mathbb{R}^m$, write $x^p = |x|^{p-1}x \in \mathbb{R}^m$. Suppose $\varphi\colon \Omega \to \mathbb{R}^m$ satisfies

$$\mathcal{L}\varphi = -\varphi^p + \xi \tag{2.1}$$

where $\xi \in \mathcal{C}^\beta(\Omega)$ with $\beta \in (-2, 0]$.

Remark 2.1. By parabolic regularity, since $\beta > -2$, (2.1) admits a unique solution $\varphi \in \mathcal{C}(\Omega)$ for any boundary data in $\mathcal{C}(\partial\Omega)$. Below we use existence (but not uniqueness) of solutions.

We derive an a priori estimate on φ that is independent of boundary data. To give some intuition for the statement and proof of the estimate, consider the *scaling critical* dimension

$$\alpha = 2/(p-1) \tag{2.2}$$

for the 'leading terms' $\mathcal{L}\varphi + \varphi^p$ in (2.1). Here α is derived from the equation

$$\alpha + 2 = \alpha p \,,$$

which represents that $\mathcal{L}\varphi$ and φ^p scale in the same way under the transformation $\varphi \mapsto \lambda^\alpha \varphi \circ T_{z,\lambda}$. Namely, if φ solves (2.1), then $\psi = \lambda^\alpha \varphi \circ T_{z,\lambda}$ solves

$$\mathcal{L}\psi = (-\lambda^{\alpha+2}\varphi^p + \lambda^{\alpha+2}\xi) \circ T_{z,\lambda} = -\psi^p + \lambda^{\alpha+2}\xi \circ T_{z,\lambda} \,, \tag{2.3}$$

where we used that $\lambda^{\alpha+2}\varphi^p \circ T_{z,\lambda} = \lambda^{\alpha+2-\alpha p}\psi^p$ and $\alpha + 2 - \alpha p = 0$. Then, denoting

$$\eta = \lambda^{\alpha+2}\xi \circ T_{z,\lambda} \in \mathcal{C}^\beta(\Omega) \tag{2.4}$$

and using (1.11), we have

$$\|\eta\|_{\mathcal{C}^\beta(\Omega)} = \lambda^{\alpha+2+\beta}\|\xi\|_{\mathcal{C}^\beta;z;\lambda} \,. \tag{2.5}$$

Hence, for $\lambda \ll 1$, ψ solves the same equation but with a small driver. Remark now that, if $\|\xi\|_{\mathcal{C}^\beta;z;\lambda} \gg 1$ and we choose $\lambda \ll 1$ such that $\|\eta\|_{\mathcal{C}^\beta(\Omega)} = r$ for some $r > 0$ small but strictly positive, then, from local stability of the equation (and ignoring for now the effects of boundary conditions), we expect $\|\psi\|_{\infty;\Omega} = \lambda^\alpha \|\varphi\|_{\infty;B_z(\lambda)} \lesssim 1$. Using (2.5), we thus expect the estimate

$$\|\varphi\|_{\infty;B_z(\lambda)} \lesssim \lambda^{-\alpha} = \lambda^{-\varrho(\alpha+2+\beta)} \sim \|\xi\|^\varrho_{\mathcal{C}^\beta;z;\lambda}$$

where

$$\varrho = \frac{\alpha}{\alpha+2+\beta} \in (0,\infty) \,. \tag{2.6}$$

The following proposition makes this heuristic precise and further shows space-time bounds independent of the boundary conditions (i.e. space-time localisation). Define the distance-to-the-boundary function $\|\!|\cdot|\!\|\colon \operatorname{cl}\Omega \to [0,1]$ by

$$\|\!|z|\!\| = \inf_{y\in\partial\Omega} |z-y|_{\mathfrak{s}} \,. \tag{2.7}$$

Proposition 2.2. *There exists $C > 0$ such that, if $\xi, \varphi \in \mathcal{C}(\operatorname{cl}\Omega)$ satisfy (2.1) in Ω, then for all $z \in \Omega$*

$$|\varphi(z)| \le C \max\{\|\xi\|^\varrho_{\mathcal{C}^\beta(\Omega)}, \|\!|z|\!\|^{-\alpha}\} \,, \tag{2.8}$$

where ϱ is given by (2.6) and where we extend ξ by 0 outside $\operatorname{cl}\Omega$ to make sense of $\|\xi\|_{\mathcal{C}^\beta(\Omega)}$.

More precisely, there exists $C > 0$ such that

$$|\varphi(z)| \le C \max\{\|\xi\|^\varrho_{\mathcal{C}^\beta;z;\mu_z}, \|\!|z|\!\|^{-\alpha}\} \,, \tag{2.9}$$

where $\mu_z = (C|\varphi(z)|^{-1/\alpha}) \wedge (\frac{1}{4}\|\!|z|\!\|)$.

Remark 2.3. With simple modifications, we can handle the equation $\mathcal{L}\varphi = -\varphi^p + \sum_{k=0}^{p-1} \xi_k \varphi^k$ where $\xi_k \in \mathcal{C}^{\beta_k}$ for $\beta_k \le 0$ satisfying scaling subcriticality for all $0 \le k \le p-1$

$$\beta_k - \alpha k > -2 - \alpha$$

and complementary Young regularity for all $1 \le k \le p-1$

$$\beta_k + 2 + \min_{0 \le i \le p-1} \beta_i > 0 \,.$$

This would cover the 'remainder' in the stochastic quantisation equation of Φ^4_2 considered by Da Prato–Debussche [DPD03], a priori estimates for which were studied in [MW17] with different methods (and without space-time localisation bounds).

We can also consider source terms with mixed homogeneity, such as $\mathcal{L}\varphi = -\varphi^p + \xi + \zeta$ with $\zeta \in \mathcal{C}^\gamma$ where $\gamma \in (-2, 0]$, and the term $\|\xi\|^\varrho_{\mathcal{C}^\beta;z;\mu_z}$ in (2.9) would become $\|\xi\|^\varrho_{\mathcal{C}^\beta;z;\mu_z} \vee \|\zeta\|^{\frac{\alpha}{\alpha+2+\gamma}}_{\mathcal{C}^\gamma;z;\mu_z}$. Taking $\gamma = 0$ recovers the main result of [MW20a].

Remark 2.4. The qualitative assumption that φ and ξ are continuous is used only to simplify the proof. We will later prove a more general result without this assumption.

Remark 2.5. There are two reasons that the 'local' quantity $\|\xi\|^\varrho_{\mathcal{C}^\beta;z;\mu_z}$ in (2.9) is more natural than the 'global' quantity $\|\xi\|^\varrho_{\mathcal{C}^\beta(\Omega)}$ in the less sharp bound (2.8).

First, since $\mu_z \le \frac{1}{4}|\!|\!|z|\!|\!|$, $\|\xi\|_{\mathcal{C}^\beta;z;\mu_z}$ depends only on the quantities $\langle \xi, \psi\rangle$ where $\operatorname{supp}\psi \subset B_z(\frac{1}{2}|\!|\!|z|\!|\!|)$ for $z \in \Omega$. In particular, $\|\xi\|_{\mathcal{C}^\beta;z;\mu_z}$ depends only on ξ as an element of $\mathcal{D}'(\Omega)$. In contrast, $\|\xi\|_{\mathcal{C}^\beta(\Omega)}$ artificially depends on ξ outside Ω.

Second, for $\beta \le \gamma \le 0$, we have the interpolation bound

$$\|\xi\|_{\mathcal{C}^\beta;z;\eta} \le \eta^{\gamma-\beta} \|\xi\|_{\mathcal{C}^\gamma;z;\eta}.$$

Suppose now we know (2.9) for some $\beta \le 0$. Then, for $\gamma \in [\beta, 0]$, either $|\varphi(z)| \lesssim |\!|\!|z|\!|\!|^{-\alpha}$, or

$$|\varphi(z)| \lesssim \|\xi\|^\varrho_{\mathcal{C}^\beta;z;\mu_z} \le \mu_z^{\varrho(\gamma-\beta)} \|\xi\|^\varrho_{\mathcal{C}^\gamma;z;\mu_z} \lesssim |\varphi(z)|^{-\varrho(\gamma-\beta)/\alpha} \|\xi\|^\varrho_{\mathcal{C}^\gamma;z;\mu_z}$$

and hence $|\varphi(z)|^{1+\varrho(\gamma-\beta)/\alpha} \lesssim \|\xi\|^\varrho_{\mathcal{C}^\gamma;z;\mu_z}$ and

$$|\varphi(z)| \lesssim \|\xi\|^{\varrho/(1+\varrho(\gamma-\beta)/\alpha)}_{\mathcal{C}^\gamma;z;\mu_z} \,.$$

Observe now that $1+\varrho(\gamma-\beta)/\alpha = 1+\frac{\gamma-\beta}{\alpha+2-\alpha i+\beta} = \frac{\alpha+2+\gamma}{\alpha+2+\beta}$ and thus $\varrho/(1-\varrho(\gamma-\beta)/\alpha) = \frac{\alpha}{\alpha+2+\gamma}$, which is the correct exponent for the bound (2.9) if we consider γ in place of β. Therefore (2.9) for a *single* choice of β implies the same bound for all $\gamma \in [\beta, 0]$ with the correct exponents. No such implication holds for the bound (2.8).

The proof of Proposition 2.2 relies on a scaling argument plus the following simple coercivity property of the equation with small drivers.

Lemma 2.6. *There exist $\varepsilon, \delta, r > 0$ such that, if $\xi \in \mathcal{D}'(\Omega^1)$ and $\varphi \in \mathcal{C}(\operatorname{cl}\Omega)$ satisfy (2.1) on Ω and $\|\xi\|_{\mathcal{C}^\beta(\Omega)} \le r$ and $\|\varphi\|_{\infty;\Omega} \le \varepsilon + \delta$, then $|\varphi(0)| < \varepsilon$.*

Proof. The proof is easy and relies on the maximum principle and a comparison to the case $\xi = 0$. We give the details as we follow the same strategy later in a more complicated setting.

Consider $\mathring{\varphi} \in \mathcal{C}(\Omega)$ solving $\mathcal{L}\mathring{\varphi} = -\mathring{\varphi}^p$ in Ω, i.e. (2.1) with $\xi = 0$. Then, for every $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) \in (0, \frac{\varepsilon}{2})$ such that, if $\|\mathring{\varphi}\|_{\infty;\Omega} \le \varepsilon + \delta$, then $|\mathring{\varphi}(0)| < \varepsilon - \delta$.

Consider now $\varphi \in \mathcal{C}(\mathrm{cl}\,\Omega)$ solving (2.1) on Ω with $\|\varphi\|_{\infty;\Omega} \le \varepsilon + \delta$. Let $\mathring{\varphi}$ be as above with boundary data $\mathring{\varphi} \restriction_{\partial\Omega} = \varphi \restriction_{\partial\Omega}$ (such $\mathring{\varphi}$ exists by parabolic regularity). Then $\|\mathring{\varphi}\|_{\infty;\Omega} \le \varepsilon + \delta$ by the maximum principle, and thus $|\mathring{\varphi}(0)| < \varepsilon - \delta$.

Furthermore, on Ω,

$$\mathcal{L}(\varphi - \mathring{\varphi}) = \mathring{\varphi}^p - \varphi^p + \xi \,.$$

Consider now a truncation K of the heat kernel such that $\mathrm{supp}(K) \subset B_0(1/2)$ and $\mathcal{L}K = \delta_0 + U$ where U is smooth and $\mathrm{supp}(U) \subset B_0(1/2)$. Define $Y \colon \bar{\Omega} \to \mathbb{R}$ by

$$\varphi - \mathring{\varphi} = K * g + Y \tag{2.10}$$

where $g = \xi + \mathbf{1}_\Omega(\mathring{\varphi}^p - \varphi^p) \in \mathcal{D}'(\Omega^1)$. Then, on Ω,

$$\mathcal{L}Y + U * g = 0 \,. \tag{2.11}$$

Since $\beta > -2$ and U is smooth with small support, we have $\|U * \xi\|_{\infty;\Omega} \lesssim \|\xi\|_{\mathcal{C}^\beta(\Omega)} \ll \delta$ whenever $\|\xi\|_{\mathcal{C}^\beta(\Omega)} \le r$ and $r \ll \delta$. Since $\|\mathring{\varphi}\|_{\infty;\Omega} \vee \|\varphi\|_{\infty;\Omega} \le 2\varepsilon$, we also have

$$\|U * \{\mathbf{1}_\Omega(\mathring{\varphi}^p - \varphi^p)\}\|_{\infty;\Omega} \lesssim \varepsilon^{p-1}\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \,.$$

It follows that

$$\|U * g\|_{\infty;\Omega} \le c\delta + C\varepsilon^{p-1}\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \tag{2.12}$$

where $c \downarrow 0$ as $r \downarrow 0$. Here and below, $C > 0$ is a constant whose value may change from line to line, but does not depend on $\varepsilon > 0$.

In a similar way, using a scale decomposition of K (see, e.g. [Hai14, Lem. 5.5]) and the fact that $(\mathring{\varphi} - \varphi) \restriction_{\partial\Omega} = 0$, we obtain for r sufficiently small

$$\|Y\|_{\infty;\partial\Omega} \le \|K * g\|_{\infty;\bar{\Omega}} \le \delta/4 + C\varepsilon^{p-1}\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \,. \tag{2.13}$$

It follows from (2.11), (2.12), (2.13), and the maximum principle that, for r sufficiently small,

$$\|Y\|_{\infty;\bar{\Omega}} \le \delta/2 + C\varepsilon^{p-1}\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \,.$$

Inserting this back into (2.10) and combining with (the second bound in) (2.13) we obtain

$$\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \le 3\delta/4 + C\varepsilon^{p-1}\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \,.$$

It follows that, for all $\varepsilon > 0$ sufficiently small, $\|\mathring{\varphi} - \varphi\|_{\infty;\Omega} \le \delta$ and therefore, since $|\mathring{\varphi}(0)| < \varepsilon - \delta$, one has $|\varphi(0)| < \varepsilon$ as desired. ∎

Proof of Proposition 2.2. Let ε, δ be as in Lemma 2.6 and $h > 1$ such that

$$1 + \varepsilon^{1/\alpha}(\varepsilon + \delta)^{-1/\alpha} h \le h \,. \tag{2.14}$$

Since we assumed that $\varphi, \xi \in \mathcal{C}(\mathrm{cl}\,\Omega)$, there exists $z \in \mathrm{cl}\,\Omega$ which maximises

$$\frac{|\varphi(z)|}{\max\{\|\xi\|^{\varrho}_{\mathcal{C}^\beta;z;h\lambda_z}, |\!|\!|z|\!|\!|^{-\alpha}\}} \stackrel{\text{def}}{=} \frac{|\varphi(z)|}{Q} \stackrel{\text{def}}{=} C_* \tag{2.15}$$

where we set $\lambda_z^\alpha = (\varepsilon|\varphi(z)|^{-1}) \wedge (\frac{1}{4h}|\!|\!|z|\!|\!|)^\alpha$. Without loss of generality, we assume $|\!|\!|z|\!|\!| > 0$. We will derive a bound on C_*, from which the proof will follow with $\mu_z = h\lambda_z$.

We first claim that there exists $C_0 > 0$ such that, if $C_* > C_0$, then

(a) $\varepsilon|\varphi(z)|^{-1} \le (\frac{1}{4h}|\!|\!|z|\!|\!|)^\alpha$, so in particular $\lambda_z^\alpha = \varepsilon|\varphi(z)|^{-1}$, and

(b) for all $x \in B_z(\lambda_z)$,

$$\|\!|x|\!\|^\alpha > \varepsilon(\varepsilon+\delta)^{-1}\|\!|z|\!\|^\alpha \,. \tag{2.16}$$

Indeed, both statements hold if $\lambda_z \ll \|\!|z|\!\|$. Since

$$\lambda_z^\alpha \le \varepsilon|\varphi(z)|^{-1} = \varepsilon C_*^{-1}Q^{-1} \le \varepsilon C_*^{-1}\|\!|z|\!\|^\alpha \,,$$

the existence of C_0 follows. We suppose henceforth that $C_* > C_0$ as otherwise (2.9) holds with $C_* = C_0$ and we are done.

Next, we define $\psi = \lambda_z^\alpha \varphi \circ T_{z,\lambda_z} \in \mathcal{C}(\mathrm{cl}\,\Omega)$ and observe that

$$|\psi(0)| = \lambda_z^\alpha|\varphi(z)| = \varepsilon \,. \tag{2.17}$$

Furthermore, by (2.3), in Ω,

$$\mathcal{L}\psi = -\psi^p + \eta \tag{2.18}$$

with $\eta \in \mathcal{C}^\beta(\Omega)$ as in (2.4). Then, by (2.5) and the definition of ϱ in (2.6),

$$\|\eta\|^\varrho_{\mathcal{C}^\beta(\Omega)} = \lambda_z^{(\alpha+2+\beta)\varrho}\|\xi\|^\varrho_{\mathcal{C}^\beta;z;\lambda_z} \le \lambda_z^\alpha Q = \varepsilon C_*^{-1} \,. \tag{2.19}$$

We next claim that $\|\psi\|_{\infty;\Omega} \le \varepsilon+\delta$, which is equivalent to

$$\sup_{x\in B_z(\lambda_z)} \varepsilon|\varphi(x)| \le (\varepsilon+\delta)|\varphi(z)| \,.$$

Arguing by contradiction, suppose there exists $x \in B_z(\lambda_z)$ such that

$$\varepsilon|\varphi(x)| > (\varepsilon+\delta)|\varphi(z)| \,. \tag{2.20}$$

Then

$$\varepsilon|\varphi(x)|^{-1} < \varepsilon^2(\varepsilon+\delta)^{-1}|\varphi(z)|^{-1} \le \varepsilon(\varepsilon+\delta)^{-1}\Big(\frac{\|\!|z|\!\|}{4h}\Big)^\alpha < \Big(\frac{\|\!|x|\!\|}{4h}\Big)^\alpha \,,$$

where we used (2.20) in the first bound, (a) in the second bound, and (2.16) in the third bound. In particular, recalling $\lambda_x^\alpha = (\varepsilon|\varphi(x)|^{-1}) \wedge (\frac{1}{4h}\|\!|x|\!\|)^\alpha$,

$$\lambda_x^\alpha = \varepsilon|\varphi(x)|^{-1} < \varepsilon^2(\varepsilon+\delta)^{-1}|\varphi(z)|^{-1} = \varepsilon(\varepsilon+\delta)^{-1}\lambda_z^\alpha \,.$$

Therefore, by (2.14), we have $\lambda_z + h\lambda_x \le h\lambda_z$, and thus $B_x(h\lambda_x) \subset B_z(h\lambda_z)$ and

$$\|\xi\|^\varrho_{\mathcal{C}^\beta;x;h\lambda_x} \le \|\xi\|^\varrho_{\mathcal{C}^\beta;z;h\lambda_z} \,. \tag{2.21}$$

It follows that

$$\begin{aligned}|\varphi(x)| &\le C_* \max\{\|\xi\|^\varrho_{\mathcal{C}^\beta;x;h\lambda_x}, \|\!|x|\!\|^{-\alpha}\} \\ &\le C_* \max\{\|\xi\|^\varrho_{\mathcal{C}^\beta;z;h\lambda_z}, \varepsilon^{-1}(\varepsilon+\delta)\|\!|z|\!\|^{-\alpha}\} \le \varepsilon^{-1}(\varepsilon+\delta)|\varphi(z)| \,,\end{aligned}$$

where we used the definition of C_* in (2.15) for the first and third bounds, and (2.16) and (2.21) for the second bound. This contradicts (2.20) and concludes the proof of $\|\psi\|_{\infty;\Omega} \le \varepsilon+\delta$.

To conclude the proof, by (2.19), there exists $C_1 > 0$ such that, if $C_* > C_1$, then $\|\eta\|_{\mathcal{C}^\beta(\Omega)} \le r$ where r is from Lemma 2.6. Since ψ solves (2.18) and $\|\psi\|_\infty \le \varepsilon+\delta$, Lemma 2.6 implies that $|\psi(0)| < \varepsilon$ if $C_* > C_1$. But this is a contradiction to (2.17), so we conclude that $C_* \le C_1$. ∎

3 Abstract bounds

We now flesh out the abstract conditions that make the above argument work. An important remark is that it is hardly important that φ solves a PDE - the important ingredients are Lemma 2.6, which controls the behaviour of the solution for small drivers, and the scaling relation (2.3). Our abstract assumptions are as follows.

Domains. Let (Σ, ϱ) be a metric space, called the 'domain', with a distinguished set $\partial\Sigma \subset \Sigma$ called the 'boundary'. We assume that, for all $z \in \Sigma$,

$$\|\!|z|\!\| \stackrel{\text{def}}{=} \inf_{x \in \partial\Sigma} \varrho(x, z) \in [0, \infty) \, .$$

We assume that, for every $z \in \Sigma$ and $0 < \lambda \leq \|\!|z|\!\|$, there are sets $\Sigma_{z,\lambda}$ with a 'boundary' $\partial\Sigma_{z,\lambda} \subset \Sigma_{z,\lambda}$ and an injection $T_{z,\lambda} \colon \Sigma_{z,\lambda} \hookrightarrow \Sigma$ such that $T_{z,\lambda}(o_{z,\lambda}) = z$ for some $o_{z,\lambda} \in \Sigma_{z,\lambda}$.

Remark 3.1. In this section, and this section only, we deviate from the notation of Section 1.1, so the map $T_{z,\lambda}$ is *not* assumed to be of the form (1.4).

Notation 3.2. Below we introduce several objects indexed by $z \in \Sigma$ and $0 < \lambda \leq \|\!|z|\!\|$. Unless otherwise stated, we let z, λ run over all such pairs.

We suppose that

$$\sup_{x \in \Sigma_{z,\lambda}} \varrho(z, T_{z,\lambda}(x)) \leq \lambda \, . \tag{3.1}$$

See Figure 1 for an illustration.

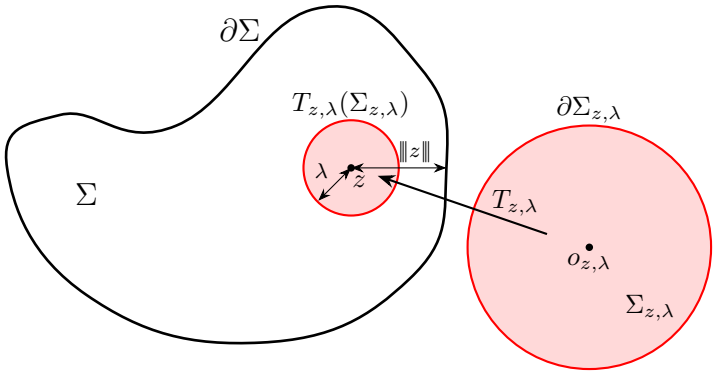


Figure 1: Example of domain configuration.

Drivers. There are sets of 'drivers' $\mathbb{M}$ and $\mathbb{M}_{z,\lambda}$, a 'scaling' map $R_{z,\lambda} \colon \mathbb{M} \to \mathbb{M}_{z,\lambda}$, and a 'norm' $\|\cdot\|_{z,\lambda} \colon \mathbb{M}_{z,\lambda} \to [0, \infty)$. We also suppose there are 'local norms' $[\![\cdot]\!]_{z,\lambda} \colon \mathbb{M} \to [0, \infty)$ satisfying the following properties. Consider $\lambda \leq \|\!|z|\!\|$, $x \in T_{z,\lambda}(\Sigma_{z,\lambda})$, and $\lambda' \leq \|\!|x|\!\|$ such that $\varrho(x, z) + \lambda' \leq \lambda$ (this condition in particular implies that the ball of radius λ' centred at x is contained in the ball of radius λ centred at z, see Figure 2). Then

$$[\![M]\!]_{x,\lambda'} \leq [\![M]\!]_{z,\lambda} \, . \tag{3.2}$$

Furthermore, we assume that there exists $\alpha > 0$ such that

$$\|R_{z,\lambda} M\|_{z,\lambda} \leq \lambda^{\alpha} [\![M]\!]_{z,\lambda} \, . \tag{3.3}$$

Remark 3.3. The maps $R_{z,\lambda}$ are the abstract analogues of $\xi \mapsto \lambda^{\alpha+2} \xi \circ T_{z,\lambda}$ from (2.4).

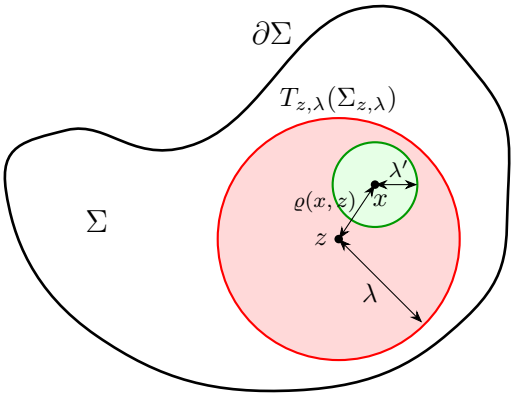


Figure 2: Example configuration of λ, λ', x, z.

Solutions. Let $(E, |\cdot|)$ be a normed vector space. We equip $E^{\Sigma_{z,\lambda}}$ with the uniform norm $\|\varphi\|_\infty = \sup_{x\in\Sigma_{z,\lambda}} |\varphi(x)|$. For $\alpha > 0$ as in (3.3), define $R_{z,\lambda}\colon E^\Sigma \to E^{\Sigma_{z,\lambda}}$ by

$$R_{z,\lambda}(\varphi) = \lambda^\alpha \varphi \circ T_{z,\lambda} \,.$$

For every $M \in \mathbb{M}$ (respectively $M \in \mathbb{M}_{z,\lambda}$) consider a subset $\mathbb{S}_M \subset E^\Sigma$ (respectively $\mathbb{S}_M \subset E^{\Sigma_{z,\lambda}}$), which we call the set of 'solutions' with respect to M. We assume that $\mathbb{S}_M$ is closed under rescaling in the sense that, for all $M \in \mathbb{M}$,

$$R_{z,\lambda}\mathbb{S}_M \subset \mathbb{S}_{R_{z,\lambda}M} \,. \tag{3.4}$$

Local coercivity. Our final assumption is that there exists $\delta > 0$ such that, for all $z \in \Sigma$, $0 < \lambda \le |\!|\!| z |\!|\!|$, $M \in \mathbb{M}_{z,\lambda}$ with $\|M\|_{z,\lambda} < 1$, and $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_\infty < 1 + \delta$, we have

$$|\varphi(o_{z,\lambda})| < 1 \,. \tag{3.5}$$

(This is the analogue of Lemma 2.6.)

Theorem 3.4. *With the above assumptions, consider $M \in \mathbb{M}$ and $\varphi \in \mathbb{S}_M$. Let*

$$\theta = (1+\delta)^{-1/\alpha} \,, \qquad \nu \ge (1-\theta)^{-1} \,.$$

Then for all $z \in \Sigma$

$$|\varphi(z)| \le \max\{[\![M]\!]_{z,\lambda_z/(1-\theta)}, \nu^\alpha |\!|\!| z |\!|\!|^{-\alpha}\} \tag{3.6}$$

where $\lambda_z = |\varphi(z)|^{-1/\alpha} \wedge (\frac{1}{\nu}|\!|\!| z |\!|\!|)$.

Before the proof, we give an immediate corollary in which we slightly generalise the form of the local coercivity assumption.

Corollary 3.5. *Suppose that, instead of local coercivity, there exist $\varepsilon, \kappa, r > 0$ such that, for all $z \in \Sigma$, $0 < \lambda \le |\!|\!| z |\!|\!|$, $M \in \mathbb{M}_{z,\lambda}$ with $\|M\|_{z,\lambda} < r$, and $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_\infty < \varepsilon$, we have $|\varphi(o_{z,\lambda})| < \varepsilon - \kappa$. Denote $1 + \delta = \frac{\varepsilon}{\varepsilon-\kappa}$ and $\theta = (1+\delta)^{-1/\alpha}$, and let $\nu \ge (1-\theta)^{-1}$. Then*

$$|\varphi(z)| \le \frac{\varepsilon}{1+\delta} \max\{r^{-1}[\![M]\!]_{z,\lambda_z/(1-\theta)}, \nu^\alpha |\!|\!| z |\!|\!|^{-\alpha}\}$$

where $\lambda_z = (\frac{1+\delta}{\varepsilon}|\varphi(z)|)^{-1/\alpha} \wedge (\frac{1}{\nu}|\!|\!| z |\!|\!|)$.

Proof. The norms $|\varphi|^\star = \frac{1+\delta}{\varepsilon}|\varphi|$ and $\|M\|^\star_{z,\lambda} = r^{-1}\|M\|_{z,\lambda}$ verify local coercivity, so the conclusion follows from Theorem 3.4. ∎

In the rest of the section, we prove Theorem 3.4. Our assumptions are designed so that the proof is similar to that of Proposition 2.2. Nonetheless, because this is one of our main results, we give the details for completeness.

Proof of Theorem 3.4. Arguing by contradiction, suppose (3.6) does not hold for some $z \in \Sigma$. Then there exists $g > \nu$ sufficiently close to ν such that

$$C_* \stackrel{\text{def}}{=} \sup_{z\in\Sigma} \frac{|\varphi(z)|}{\max\{[\![M]\!]_{z,\lambda_z/(1-\theta)}, g^\alpha \|z\|^{-\alpha}\}} > 1 \,. \tag{3.7}$$

Let $z \in \Sigma$ such that

$$|\varphi(z)| = \bar{C}Q \,, \quad Q = \max\{[\![M]\!]_{z,\lambda_z/(1-\theta)}, g^\alpha \|z\|^{-\alpha}\} \tag{3.8}$$

where $\bar{C} \in (1, C_*]$ is sufficiently close to C_* such that

$$C_*/\bar{C} < (1-g^{-1})^\alpha(1+\delta) \,. \tag{3.9}$$

(Note that the right-hand side of (3.9) is strictly larger than 1 since $g > \nu \geq (1-\theta)^{-1}$, thus such z and $\bar{C}$ exist.) Then, using (3.8), $\bar{C} > 1$, and $Q^{-1/\alpha} \leq g^{-1}\|z\|$, we have

$$\lambda_z \leq |\varphi(z)|^{-1/\alpha} = (\bar{C}Q)^{-1/\alpha} < g^{-1}\|z\| \,. \tag{3.10}$$

In particular, $|\varphi(z)|^{-1/\alpha} < \nu^{-1}\|z\|$ and therefore $\lambda_z^\alpha = |\varphi(z)|^{-1}$.

Next, denote $\psi = R_{z,\lambda_z}\varphi = \lambda_z^\alpha \varphi \circ T_{z,\lambda_z}$. Remark that $\psi \in \mathbb{S}_{R_{z,\lambda_z}M}$ by (3.4). Furthermore

$$|\psi(o_{z,\lambda_z})| = \lambda_z^\alpha |\varphi(z)| = 1 \,. \tag{3.11}$$

Note that, by (3.3),

$$\|R_{z,\lambda_z}M\|_{z,\lambda_z} \leq \lambda_z^\alpha Q = \bar{C}^{-1} < 1 \,. \tag{3.12}$$

We now claim that $\|\psi\|_\infty < 1+\delta$, which is equivalent to

$$\sup_{x\in T_{z,\lambda_z}(\Sigma_{z,\lambda_z})} |\varphi(x)| < (1+\delta)|\varphi(z)| \,.$$

Suppose for sake of contradiction that there exists $x \in T_{z,\lambda_z}(\Sigma_{z,\lambda_z})$ such that

$$|\varphi(x)| \geq (1+\delta)|\varphi(z)| \,. \tag{3.13}$$

By (3.10), $\lambda_z < g^{-1}\|z\| < \|z\|$. Furthermore, $\varrho(z,x) \leq \lambda_z$ by (3.1). Therefore

$$\|x\| \geq \|z\| - \varrho(x,z) \geq \|z\| - \lambda_z > (1-g^{-1})\|z\| \,. \tag{3.14}$$

Then, recalling $\theta = (1+\delta)^{-1/\alpha}$,

$$|\varphi(x)|^{-1/\alpha} \leq \theta|\varphi(z)|^{-1/\alpha} < \theta g^{-1}\|z\| < g^{-1}\|x\| < \nu^{-1}\|x\|$$

where we used (3.13) in the first bound, (3.10) in the second bound, and (3.14) and the fact that $\theta < 1-g^{-1}$ due to $g > \nu \geq (1-\theta)^{-1}$ in the third bound. Hence

$$\lambda_x = |\varphi(x)|^{-1/\alpha} \leq (1+\delta)^{-1/\alpha}|\varphi(z)|^{-1/\alpha} = \theta\lambda_z \tag{3.15}$$

and thus

$$\varrho(x,z) + \lambda_x/(1-\theta) \le \lambda_z + \lambda_x/(1-\theta) \le \lambda_z/(1-\theta)$$

where we used $\varrho(x,z) \le \lambda_z$ in the first bound and (3.15) in the second bound. Therefore, applying (3.2) with $\lambda' = \lambda_x/(1-\theta)$ and $\lambda = \lambda_z/(1-\theta)$, we obtain

$$[\![M]\!]_{x,\lambda_x/(1-\theta)} \le [\![M]\!]_{z,\lambda_z/(1-\theta)} \,. \tag{3.16}$$

It follows that

$$\begin{aligned} |\varphi(x)| &\le C_* \max\{[\![M]\!]_{x,\lambda_x/(1-\theta)}, g^\alpha |\!|\!| x |\!|\!|^{-\alpha}\} \\ &\le C_* \max\{[\![M]\!]_{z,\lambda_z/(1-\theta)}, g^\alpha |\!|\!| z |\!|\!|^{-\alpha}(1-g^{-1})^{-\alpha}\} \le (1-g^{-1})^{-\alpha} C_* Q \,, \end{aligned} \tag{3.17}$$

where in the first inequality we used the definition of C_* in (3.7), and in the second inequality we used (3.14) and (3.16). However, by (3.8)-(3.9),

$$(1-g^{-1})^{-\alpha} C_* Q = (1-g^{-1})^{-\alpha} C_* |\varphi(z)| \bar{C}^{-1} < |\varphi(z)|(1+\delta) \,,$$

which, together with (3.17), contradicts (3.13). We have thus proved the claim $\|\psi\|_\infty < 1+\delta$.

Since $\psi \in \mathbb{S}_{R_{z,\lambda_z} M}$ with $\|\psi\|_\infty < 1+\delta$ and $\|R_{z,\lambda_z} M\|_{z,\lambda_z} < 1$ by (3.12), it follows from the coercivity assumption (3.5) that $|\psi(o_{z,\lambda})| < 1$, which contradicts (3.11). ∎

4 Applications to well-posed equations

The purpose of this section is to give examples of how Theorem 3.4 implies a priori estimates for classically well-posed equations.

4.1 Classical case revisited

As a first example, we show how Proposition 2.2 follows from Theorem 3.4.

Domains. We take our domain as $\Sigma = \operatorname{cl}\Omega$ with boundary $\partial\Omega \subset \operatorname{cl}\Omega$ as in Section 1.1.2. We take the metric on $\operatorname{cl}\Omega$ as $\varrho(x,y) = |x-y|_{\mathfrak{s}}$, i.e. the parabolic distance, so that the 'distance to the boundary' $|\!|\!|\cdot|\!|\!|\colon \operatorname{cl}\Omega \to [0,1]$ is $|\!|\!| z |\!|\!| = \inf_{y\in\partial\Omega} \varrho(z,y)$ as in (2.7).

For every $z \in \Omega$ and $0 < \lambda \le |\!|\!| z |\!|\!|$, we take $\Sigma_{z,\lambda} = \Omega$ and $T_{z,\lambda}\colon \Omega \to \operatorname{cl}\Omega$ as defined by (1.4). (In particular, $o_{z,\lambda} = 0$.) Note that these choices clearly satisfy (3.1).

Drivers. Our spaces of drivers are $\mathbb{M} = \mathbb{M}_{z,\lambda} = \mathcal{C}^\beta(\Omega)$ with norm

$$\|\xi\|_{z,\lambda} = \|\xi\|^\varrho_{\mathcal{C}^\beta(\Omega)}$$

where $\varrho = \frac{\alpha}{2+\alpha+\beta}$ as in (2.6) with $\alpha = 2/(p-1)$ as in (2.2). The rescaling operators are

$$R_{z,\lambda}\xi = \lambda^{2+\alpha} \xi \circ T_{z,\lambda} \,.$$

The local norms are

$$[\![\xi]\!]_{z,\lambda} = \|\xi\|^\varrho_{\mathcal{C}^\beta;z;\lambda}$$

for which (3.2) clearly holds and (3.3) follows from

$$\|\lambda^{\alpha+2}\xi \circ T_{z,\lambda}\|^\varrho_{\mathcal{C}^\beta(\Omega)} = \lambda^{(\alpha+2+\beta)\varrho} \|\xi\|^\varrho_{\mathcal{C}^\beta;z;\lambda} = \lambda^\alpha \|\xi\|^\varrho_{\mathcal{C}^\beta;z;\lambda} \,,$$

where we used (2.5) (or (2.19)).

Solutions. We set $E = \mathbb{R}^m$ as the target space and let $\mathbb{S}_\xi$ be the set of functions $\varphi \in \mathcal{C}(\operatorname{cl}\Omega)$ that satisfy (2.1) in Ω. The closure under rescaling assumption (3.4) is verified in (2.3).

Local coercivity, in the form of Corollary 3.5, follows from Lemma 2.6.

We therefore satisfy all the conditions of Corollary 3.5 and conclude that there exist $C > 0$ and $\theta \in (0,1)$ such that, for all $\nu \geq (1-\theta)^{-1}$,

$$|\varphi(z)| \leq C \max\{[\![\xi]\!]_{z,\lambda_z/(1-\theta)}, \nu^\alpha |\!|\!| z |\!|\!|^{-\alpha}\} ,$$

where $\lambda_z = |\varphi(z)|^{-1/\alpha} \wedge (\frac{1}{\nu}|\!|\!| z |\!|\!|)$. Taking $\nu = 4(1-\theta)^{-1}$ recovers Proposition 2.2.

4.2 Young ODEs

We consider a Young ODE

$$\mathrm{d}\varphi_t = -\varphi_t^p \,\mathrm{d}t + f(\varphi_t)\,\mathrm{d}X_t , \tag{4.1}$$

where $p \in (1,\infty)$ and we denote $\varphi^p = |\varphi|^{p-1}\varphi$, $X \in \mathcal{C}^\gamma([-1,0],\mathbb{R}^n)$, $\varphi \in \mathcal{C}^\gamma([-1,0],\mathbb{R}^m)$, $f \in \mathcal{C}^\theta(\mathbb{R}^m, L(\mathbb{R}^n,\mathbb{R}^m))$, and $\gamma \in (\frac{1}{2},1]$, $\theta \in (\frac{1}{\gamma}-1, 1]$.

Note that the condition $f \in \mathcal{C}^\theta$ implies that $f(\varphi) \in \mathcal{C}^{\theta\gamma}$ where $\gamma + \theta\gamma > 1$, thus $\int_s^t f(\varphi_u)\,\mathrm{d}X_u$ is well-defined as a Young integral in the sense of [FH20, Sec. 4]. The equation (4.1) is then understood in the integral sense

$$\varphi_t = \varphi_s - \int_s^t \varphi_u^p \,\mathrm{d}u + \int_s^t f(\varphi_u)\,\mathrm{d}X_u .$$

Such equations with $\gamma \in (\frac{1}{2},1]$ were studied in [Lyo94]. In this subsection, we set

$$\alpha = 1/(p-1) .$$

For $\beta \in [0,1]$, we define the β-Hölder semi-norm

$$[\varphi]_\beta = \sup_{s\neq t} |t-s|^{-\beta}|\varphi_t - \varphi_s|$$

and likewise for $[f]_\theta$. We also define

$$\|f\|_{\mathcal{C}^\theta} = [f]_\theta + \|f\|_\infty$$

and the restricted β-Hölder semi-norm

$$[\varphi]_{\beta;\leq h} = \sup_{0<|t-s|\leq h} |t-s|^{-\beta}|\varphi_t - \varphi_s| .$$

We also write $\|\varphi\|_{\infty;I}$, $[\varphi]_{\beta;I}$, etc. when we restrict the variables s,t to a domain I. For future reference, we record Young's estimate for $\beta,\beta' \in (0,1]$ with $\beta+\beta' > 1$, see e.g. [FH20, Sec. 4],

$$\Big|\int_s^t Y_u\,\mathrm{d}X_u - Y_s(X_t - X_s)\Big| \lesssim |t-s|^{\beta+\beta'}[X]_{\beta;[s,t]}[Y]_{\beta';[s,t]} . \tag{4.2}$$

In what follows, we will derive several a priori estimates on the solution φ to (4.1) using Theorem 3.4. We first place ourselves in the setting of Section 3.

Our **domains** are just $\Sigma = \Sigma_{z,\lambda} = [-1,0]$ with metric $\varrho(x,y) = |x-y|$ and rescaling maps

$$T_{z,\lambda}\colon [-1,0] \to [-1,0] , \quad T_{z,\lambda}(x) = z + \lambda x .$$

Note that $o_{z,\lambda} = 0$.

Our spaces of **drivers** $\mathbb{M} = \mathbb{M}_{z,\lambda}$ are now all pairs $M = (f, X)$ with $f \in \mathcal{C}^\theta$ and $X \in \mathcal{C}^\gamma$ defined on the appropriate spaces. Define $\psi = \lambda^\alpha \varphi \circ T_{z,\lambda}$. Then

$$\mathrm{d}\psi_t = \lambda^{1+\alpha}(\mathrm{d}\varphi)_{T_{z,\lambda}(t)} = -\lambda^{1+\alpha-\alpha p}\psi_t^p \,\mathrm{d}t + f(\lambda^{-\alpha}\psi)\,\mathrm{d}Y_t = -\psi_t^p \,\mathrm{d}t + f_\lambda(\psi)\,\mathrm{d}Y_t \,, \tag{4.3}$$

where $f_\lambda = f(\lambda^{-\alpha}\cdot)$ and $\mathrm{d}Y_t = \lambda^{1+\alpha}(\mathrm{d}X)_{T_{z,\lambda}(t)}$, i.e. $Y_t = \lambda^\alpha X_{T_{z,\lambda}(t)}$. This motivates the rescaling map

$$R_{z,\lambda}(f, X) = (f_\lambda, \lambda^\alpha X \circ T_{z,\lambda}) \,.$$

Note that $[f_\lambda]_\theta = \lambda^{-\alpha\theta}[f]_\theta$ and thus

$$\|f_\lambda\|_{\mathcal{C}^\theta} \le \lambda^{-\alpha\theta}\|f\|_{\mathcal{C}^\theta} \tag{4.4}$$

and furthermore

$$[X \circ T_{z,\lambda}]_{\gamma;[-1,0]} = \lambda^\gamma [X]_{\gamma;[z-\lambda,z]} \,, \tag{4.5}$$

The **solution** space $\mathbb{S}_M$ for $M = (f, X)$ is the set of all solutions to (4.1). The closure under rescaling assumption (3.4) is verified in (4.3).

To apply Theorem 3.4 (or rather Corollary 3.5), it remains to verify the **local coercivity** condition for suitable norms $[\![M]\!]_{z,\lambda}$ and $\|M\|_{z,\lambda}$ satisfying (3.2)-(3.3). In the rest of the section, we progressively derive sharper and more general bounds by making different choices of norms.

4.2.1 Simple bound

We first derive a simple but non-sharp bound. Fix $\varepsilon > 0$ and suppose $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_\infty \le \varepsilon$. Consider also $\mathring{\varphi}$ solving (4.1) with $X = 0$ and with the same initial condition $\mathring{\varphi}_{-1} = \varphi_{-1}$. Then there exists $\delta > 0$ such that $|\mathring{\varphi}_0| < \varepsilon - 2\delta$. Moreover $\varphi - \mathring{\varphi}$ solves

$$\mathrm{d}(\varphi - \mathring{\varphi})_t = -(\varphi^p - \mathring{\varphi}^p)_t \,\mathrm{d}t + f(\varphi_t)\,\mathrm{d}X_t \,.$$

Therefore, provided that $\varepsilon > 0$ is sufficiently small, $\|\varphi - \mathring{\varphi}\|_\infty \lesssim \|\int_{-1}^{\cdot} f(\varphi_t)\,\mathrm{d}X_t\|_\infty$. In turn,

$$\Big\| \int_{-1}^{\cdot} f(\varphi_t)\,\mathrm{d}X_t \Big\|_\infty \lesssim \|f\|_\infty [X]_\gamma + [f(\varphi)]_{\theta\gamma}[X]_\gamma \,.$$

Furthermore, $[f(\varphi)]_{\theta\gamma} \le [f]_\theta [\varphi]_\gamma^\theta$ and, by Young's estimate (4.2),

$$|\varphi_t - \varphi_s| \lesssim \varepsilon^p |t-s| + |t-s|^\gamma \|f\|_\infty [X]_\gamma + |t-s|^{\gamma+\gamma\theta}[\varphi]_\gamma^\theta [f]_\theta [X]_\gamma \,.$$

Provided $\|f\|_{\mathcal{C}^\theta}[X]_\gamma$ is small, it follows that $[\varphi]_\gamma \lesssim 1$ and thus

$$\|\varphi - \mathring{\varphi}\|_\infty \lesssim \|f\|_{\mathcal{C}^\theta}[X]_\gamma \,.$$

This motivates the norms

$$[\![(f, X)]\!]_{z,\lambda} = \big(\|f\|_{\mathcal{C}^\theta}[X]_{\gamma;[z-\lambda,z]}\big)^\varrho \,, \qquad \|(f, X)\|_{z,\lambda} = \big(\|f\|_{\mathcal{C}^\theta}[X]_\gamma\big)^\varrho \,,$$

where

$$\varrho = \frac{\alpha}{\alpha + \gamma - \alpha\theta} \,.$$

Then, by (4.4)-(4.5),

$$\|R_{z,\lambda}(f, X)\|_{z,\lambda} = \lambda^{(\alpha+\gamma-\alpha\theta)\varrho}\big(\|f\|_{\mathcal{C}^\theta}[X]_{\gamma;[z-\lambda,z]}\big)^\varrho \le \lambda^\alpha [\![(f, X)]\!]_{z,\lambda} \,,$$

which verifies (3.3). Moreover (3.2) clearly holds.

In conclusion, there exists $r > 0$ such that, if $\|(f, X)\|_{z,\lambda} \le r$, then $\|\varphi - \mathring{\varphi}\|_\infty \le \delta$ and thus $|\varphi_0| < \varepsilon - \delta$. This verifies local coercivity in the form of Corollary 3.5. It follows that there exists $C > 0$ such that, for all $z \in (-1, 0]$,

$$|\varphi_z| \le C \max\{[\![(f, X)]\!]_{z,\lambda_z}, |z+1|^{-\alpha}\} \,, \tag{4.6}$$

where $\lambda_z = (C|\varphi_z|^{-1/\alpha}) \wedge \frac{1}{2}|z+1|$.

4.2.2 Sharper bounds

We next show how, with localised Hölder norms, we can sharpen the bound (4.6). Fix again $\varepsilon > 0$ and let $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_\infty \le \varepsilon$. Consider $\mathring{\varphi}$ as at the start of Section 4.2.1. Since

$$\|\varphi - \mathring{\varphi}\|_\infty \lesssim \Big\| \int_{-1}^{\cdot} f(\varphi_t)\,\mathrm{d}X_t \Big\|_\infty \,, \tag{4.7}$$

it suffices to bound the right-hand side.

By Young's estimate (4.2), for all $h \in (0,1]$ and $|t-s| \le h$ we have

$$\Big| \int_s^t f(\varphi_u)\,\mathrm{d}X_u - f(\varphi_s)(X_t - X_s) \Big| \lesssim |t-s|^{\gamma+\gamma\theta} [f(\varphi)]_{\theta\gamma;\le h} [X]_\gamma \,. \tag{4.8}$$

By the triangle inequality, for any $t \in [-1,0]$ and $h \in (0,1]$,

$$\Big| \int_{-1}^t f(\varphi_t)\,\mathrm{d}X_t \Big| \le \sum_{0 \le i < (t+1)h^{-1}} \Big| \int_{ih-1}^{t \wedge ((i+1)h-1)} f(\varphi_s)\,\mathrm{d}X_s \Big| \,. \tag{4.9}$$

Therefore, applying (4.8) to every term on the right-hand side, we obtain

$$\Big\| \int_{-1}^{\cdot} f(\varphi_t)\,\mathrm{d}X_t \Big\|_\infty \lesssim h^{-1}(W_h + V_h) \,, \tag{4.10}$$

where

$$W_h = h^\gamma \|f\|_\infty [X]_\gamma \,, \qquad V_h = h^{\gamma+\gamma\theta} [f(\varphi)]_{\theta\gamma;\le h} [X]_\gamma \,.$$

To bound V_h and $[f(\varphi)]_{\theta\gamma;\le h}$, note that, for $|s-t| \le h$,

$$|\varphi_t - \varphi_s| \lesssim \varepsilon^p |t-s| + |t-s|^\gamma h^{-\gamma} W_h + |t-s|^{\gamma+\gamma\theta} h^{-\gamma-\gamma\theta} V_h \,,$$

from which it follows that

$$[f(\varphi)]_{\theta\gamma;\le h} \le [f]_\theta [\varphi]^\theta_{\gamma;\le h} \lesssim [f]_\theta (h^{1-\gamma} + h^{-\gamma} W_h + h^{-\gamma} V_h)^\theta = [f]_\theta (h + W_h + V_h)^\theta h^{-\gamma\theta} \,.$$

Therefore

$$V_h \lesssim h^\gamma [X]_\gamma [f]_\theta (h + W_h + V_h)^\theta \,.$$

Provided $h^\gamma [X]_\gamma [f]_\theta \ll \{h^\gamma [X]_\gamma [f]_\theta (h + W_h)^\theta\}^{1-\theta}$, i.e.

$$\Gamma \stackrel{\text{def}}{=} (h + W_h)^{\theta-1} h^\gamma [X]_\gamma [f]_\theta \ll 1 \,,$$

we obtain by a continuity argument in h that

$$V_h \lesssim h^\gamma [X]_\gamma [f]_\theta (h + W_h)^\theta = (h + W_h)\Gamma \,.$$

(If $\theta = 1$, then we do not need a continuity argument.)

We now choose h so as to make the contributions of h and W_h comparable. Specifically, suppose that $\|f\|_\infty [X]_\gamma \le \kappa$ for some small $\kappa > 0$. Choose the smallest $h \in [0,1]$ such that $h^{-1} W_h \le \kappa$. For $h = 0$, we understand $h^{-1} W_h$ as the limit $h \downarrow 0$, which exists by monotonicity. Remark that $h = 0$ if and only if $\|f\|_\infty [X]_\gamma = 0$ or if $\gamma = 1$.

If $\gamma = 1$, then $\| \int_{-1}^{\cdot} f(\varphi_t)\,\mathrm{d}X_t \|_\infty \lesssim \kappa$ due to (4.10) and the fact that $\lim_{h \downarrow 0} h^{-1} V_h = 0$.

If $\gamma \in (\frac{1}{2}, 1)$, then

$$h^{-1} V_h \lesssim \Gamma \lesssim h^{\gamma+\theta-1} [f]_\theta [X]_\gamma = (\|f\|_\infty [X]_\gamma)^{\frac{\gamma+\theta-1}{1-\gamma}} [f]_\theta [X]_\gamma \le (\|f\|_\infty^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta}^{\frac{1-\gamma}{\theta}} [X]_\gamma)^{\frac{\theta}{1-\gamma}} \,. \tag{4.11}$$

For all $\gamma \in (\frac{1}{2}, 1]$, this motivates us to (re)define the norms

$$[\![(f,X)]\!]_{z,\lambda} = \big(\|f\|_\infty^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]}\big)^\varrho \,, \qquad \|(f,X)\|_{z,\lambda} = \big(\|f\|_\infty^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta}^{\frac{1-\gamma}{\theta}} [X]_\gamma\big)^\varrho \,,$$

where

$$\varrho = \frac{\alpha}{\gamma(1+\alpha)} \,. \tag{4.12}$$

(Note that we use the *norm* $\|f\|_{\mathcal{C}^\theta}$ and not the *semi-norm* $[f]_\theta$, so smallness of $\|(f,X)\|_{z,\lambda}$ ensures smallness of $\|f\|_\infty [X]_\gamma$ that we required just above (4.11).) Then, by (4.4)-(4.5),

$$\begin{aligned}
\|R_{z,\lambda}(f,X)\|_{z,\lambda} &= \lambda^{(\alpha+\gamma)\varrho} \Big(\|f_\lambda\|_\infty^{1-\frac{1-\gamma}{\theta}} \|f_\lambda\|_{\mathcal{C}^\theta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]} \Big)^\varrho \\
&\le \lambda^{(\alpha+\gamma)\varrho - \frac{1-\gamma}{\theta}\alpha\theta\varrho} \Big(\|f\|_\infty^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]} \Big)^\varrho \\
&= \lambda^\alpha [\![(f,X)]\!]_{z,\lambda} \,,
\end{aligned}$$

which verifies (3.3). Moreover (3.2) clearly holds.

In conclusion, taking $\kappa > 0$ above sufficiently small, we can choose $r > 0$ such that, if $\|(f,X)\|_{z,\lambda} \le r$, then, by (4.7) and (4.10), $\|\varphi - \mathring{\varphi}\|_\infty \le \delta$ and thus $|\varphi_0| < \varepsilon - \delta$. This implies the local coercivity condition in the sense of Corollary 3.5. It follows that, there exists $C > 0$ such that, for all $z \in (-1, 0]$,

$$|\varphi_z| \le C \max\{ [\![(f,X)]\!]_{z,\lambda_z}, |z+1|^{-\alpha} \} \,, \tag{4.13}$$

where $\lambda_z = (C|\varphi_z|^{-1/\alpha}) \wedge \frac{1}{2}|z+1|$. This bound improves (4.6).

Remark 4.1. [BCMW22, Thm. 1.4, (1.4)] in the Young regime is a special case of (4.13).

4.2.3 Example of weighted norm

We now consider a simple example of a weighted norm on f. In Section 5 we will generalise this example (at least for $\theta = 1$) by treating more general weights and rough differential equations.

We fix $\eta \in \mathbb{R}$ and, as earlier, $\theta \in (\frac{1}{\gamma} - 1, 1]$. For $x \in \mathbb{R}^m$, define

$$|x|_\eta = \begin{cases} |x|^\eta + 1 & \text{if } \eta \ge 0 \,, \\ |x|^\eta & \text{if } \eta < 0 \,. \end{cases} \tag{4.14}$$

Note that, for $r \ge 1$, $|rx|_\eta \le r^\eta |x|_\eta$. We then define the weighted norms

$$\begin{aligned}
&\|f\|_{\infty;\eta} = \sup_{x \in \mathbb{R}^m} \frac{|f(x)|}{|x|_\eta} \,, \qquad [f]_{\theta;\eta} = \sup_{x \ne y,\, \frac{1}{2}|x| \le |y| \le |x|} \frac{|f(x) - f(y)|}{|x-y|^\theta |x|_{\eta-\theta}} \,, \\
&\|f\|_{\mathcal{C}^\theta;\eta} = \|f\|_{\infty;\eta} + [f]_{\theta;\eta} \,.
\end{aligned} \tag{4.15}$$

Note that, for $\eta < 0$ and $\|f\|_{\infty;\eta} < \infty$, one has $|f(x)| \to 0$ as $|x| \to \infty$ while $|f(x)|$ may blow up as $x \to 0$. A similar remark applies to $[f]_{\theta;\eta}$ for $\eta < \theta$.

Lemma 4.2. *For all $\eta \in \mathbb{R}$ and $\lambda \in (0,1]$, $\|f_\lambda\|_{\infty;\eta} \le \lambda^{-\alpha\eta} \|f\|_{\infty;\eta}$ and $[f_\lambda]_{\theta;\eta} \le \lambda^{-\alpha\eta} [f]_{\theta;\eta}$.*

Proof. Since $|\lambda^{-\alpha} x|_\eta \le \lambda^{-\alpha\eta} |x|_\eta$, one has

$$|f_\lambda(x)| = |f(\lambda^{-\alpha} x)| \le \lambda^{-\alpha\eta} |x|_\eta \|f\|_{\infty;\eta} \,,$$

hence $\|f_\lambda\|_{\infty;\eta} \le \lambda^{-\alpha\eta} \|f\|_{\infty;\eta}$. Furthermore, for $\frac{1}{2}|x| \le |y| \le |x|$,

$$|f_\lambda(x) - f_\lambda(y)| = |f(\lambda^{-\alpha} x) - f(\lambda^{-\alpha} y)| \le \lambda^{-\alpha\theta} |x-y|^\theta \lambda^{-\alpha(\eta-\theta)} |x|_{\eta-\theta} [f]_{\theta;\eta} \,,$$

hence $[f_\lambda]_{\theta;\eta} \le \lambda^{-\alpha\eta} [f]_{\theta;\eta}$. ∎

Suppose now that $\alpha + \gamma - \alpha\eta > 0$. We then (re)define the norms

$$[\![(f,X)]\!]_{z,\lambda} = \big(\|f\|_{\infty;\eta}^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta;\eta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]}\big)^\varrho \,, \qquad \|(f,X)\|_{z,\lambda} = \big(\|f\|_{\infty;\eta}^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta;\eta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma}\big)^\varrho \,,$$

where

$$\varrho = \frac{\alpha}{\alpha + \gamma - \alpha\eta} \,.$$

Then, by Lemma 4.2,

$$\begin{aligned}
\|R_{z,\lambda}(f,X)\|_{z,\lambda} &= \lambda^{(\alpha+\gamma)\varrho} \big(\|f_\lambda\|_{\infty;\eta}^{1-\frac{1-\gamma}{\theta}} \|f_\lambda\|_{\mathcal{C}^\theta;\eta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]}\big)^\varrho \\
&\le \lambda^{(\alpha+\gamma)\varrho - \alpha\eta\varrho} \big(\|f\|_{\infty;\eta}^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta;\eta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]}\big)^\varrho \\
&= \lambda^\alpha [\![(f,X)]\!]_{z,\lambda} \,,
\end{aligned}$$

which verifies (3.3).

It remains to verify the local coercivity condition in the form of Corollary 3.5. As in Section 4.2.2, let φ be a solution with $\|\varphi\|_\infty \le \varepsilon$. If $|\varphi_0| \le \varepsilon/2$, then we are done.

If $|\varphi_t| \in [\frac{\varepsilon}{2}, \varepsilon]$ for all $t \in [-1, 0]$, then because the weighted norms of f on the subsets $\{y : |y| \in [\frac{\varepsilon}{2}, \varepsilon]\}$ are, up to a multiple depending on ε, equivalent to the unweighted norms, the argument in Section 4.2.2 proves that $|\varphi_0| < \varepsilon - \delta$ for some $\delta > 0$ whenever $\|(f,X)\|_{z,\lambda} \le r$ for some $r \ll 1$, which verifies local coercivity in this case.

Finally, if $|\varphi_0| > \varepsilon/2$ and $|\varphi_t| < \varepsilon/2$ for some $t \in [-1, 0]$, then consider $I = [b, 0]$ where $b = \sup\{t \in [-1, 0] : |\varphi_t| < \varepsilon/2\}$, so that $I \subset [-1, 0]$ is the longest interval containing 0 for which $|\varphi_t| \ge \varepsilon/2$ for all $t \in I$. Similarly to Section 4.2.2, consider the solution $\mathring{\varphi} \colon I \to \mathbb{R}^m$ to (4.1) with initial condition $\mathring{\varphi}_b = \varphi_b$ and $X = 0$. Note that $|\mathring{\varphi}_b| = \varepsilon/2$ thus $|\mathring{\varphi}_0| \le \frac{\varepsilon}{2}$. Again by equivalence of weighted and unweighted norms of f on $\{y : |y| \in [\frac{\varepsilon}{2}, \varepsilon]\}$, if $\|(f,X)\|_{z,\lambda} \le r$ for $r \ll 1$, then $\|\varphi - \mathring{\varphi}\|_{\infty;I} < \delta$ for some $\delta \ll \varepsilon$, which finishes the proof of local coercivity.

It follows from Corollary 3.5 that there exists $C > 0$ such that, for all $z \in (-1, 0]$,

$$|\varphi_z| \le C \max\{[\![(f,X)]\!]_{z,\lambda_z}, |z+1|^{-\alpha}\} \,, \tag{4.16}$$

where $\lambda_z = (C|\varphi_z|^{-1/\alpha}) \wedge \frac{1}{2}|z+1|$.

Remark 4.3. Recalling that $\alpha = \frac{1}{p-1}$, we have $\varrho = \frac{1}{1+(p-1)\gamma-\eta}$ which matches the exponent in [BCMW22, Thm 1.4, (1.6)]. Hence, in the Young regime, (4.16) generalises the latter by giving precise dependence on f, lowering the regularity of f from $\mathcal{C}^1$ to $\mathcal{C}^\theta$, and allowing $\eta < 1$.

Remark 4.4. The norms (4.15) are modelled by the behaviour of polynomials and we chose to consider them for simplicity. However, the method applies to other norms. For example, for $\eta \in [0, \theta)$, it is arguably more natural to use the smaller norm $\|f\|_{\mathcal{C}^\theta;\theta}$ in place of $\|f\|_{\mathcal{C}^\theta;\eta}$. Then the same argument as earlier shows

$$|\varphi_z| \lesssim \max\{[\![(f,X)]\!]'_{z,\lambda_z}, |z+1|^{-\alpha}\} \,,$$

where

$$[\![(f,X)]\!]'_{z,\lambda} = \Big(\|f\|_{\infty;\eta}^{1-\frac{1-\gamma}{\theta}} \|f\|_{\mathcal{C}^\theta;\theta}^{\frac{1-\gamma}{\theta}} [X]_{\gamma;[z-\lambda,z]}\Big)^\varrho$$

and

$$\varrho = \frac{\alpha}{\gamma + \alpha(\gamma - (1 - \frac{1-\gamma}{\theta})\eta)} \,,$$

which sharpens (4.12) in the case that $\eta = 0$.

5 Rough paths

We now generalise Section 4.2 by considering a *rough differential equation* (RDE)

$$\mathrm{d}\varphi_t = -\varphi_t^p \,\mathrm{d}t + f(\varphi_t)\,\mathrm{d}\mathbf{X}_t \,, \tag{5.1}$$

where $\varphi \in \mathcal{C}^\gamma([-1,0],\mathbb{R}^m)$, $p \in (1,\infty)$ and we write $\varphi^p = |\varphi|^{p-1}\varphi$, $\mathbf{X}$ is an n-dimensional γ-Hölder branched rough path for $\gamma \in (0,1]$ in the sense of [Gub10] (see also [HK15, BC19, BCFP19, FH20, BCMW22]), and $f \in \mathcal{C}^N(\mathbb{R}^m, L(\mathbb{R}^n,\mathbb{R}^m))$ where $N = \lfloor \frac{1}{\gamma} \rfloor$. (In this regime, one in general has existence but not uniqueness of solutions to (5.1).)

Our main result in this section is Theorem 5.5, which establishes a priori bounds on φ under very general (multi-linear) conditions on f. In Section 5.2, we show a consequence of Theorem 5.5 under a simple (linear) condition on f, which strengthens the main result of [BCMW22].

Remark 5.1. It is not crucial for our argument that the drift is polynomial: by including the drift suitably into the space of 'drivers', we can handle any sufficiently regular function that behaves like $-\varphi^p$ for large $|\varphi|$.

Remark 5.2. We restrict to $f \in \mathcal{C}^N$ for simplicity, but with more careful arguments, similar to those in Section 4.2, one can handle the general case $f \in \mathcal{C}^\theta$ where $\theta \in (\frac{1}{\gamma} - 1, N]$.

5.1 General a priori estimates

Recall that $\mathbf{X}$ is indexed by (labelled rooted) forests τ with number of vertices $|\tau| \le N$ and with every vertex labelled by an element of the set $[n]$, and $\mathbf{X}^\tau \colon [-1,0]^2 \to \mathbb{R}$ is continuous with

$$\|\mathbf{X}^\tau\| \stackrel{\text{def}}{=} \|\mathbf{X}^\tau\|_{\gamma|\tau|} \stackrel{\text{def}}{=} \sup_{s\neq t} \frac{|\mathbf{X}^\tau_{s,t}|}{|t-s|^{\gamma|\tau|}} < \infty \,.$$

We denote by $\mathbf{1}$ the empty forest for which $\mathbf{X}^{\mathbf{1}} \equiv 1$. For an interval $I \subset [-1,0]$, we write $\|\mathbf{X}^\tau\|_I$ for the corresponding semi-norm obtained by restricting to $s,t \in I$.

Recall that every forest can be written uniquely as $\tau = [\sigma_1]_{i_1} \cdots [\sigma_\ell]_{i_\ell}$, where $i_j \in [n]$, σ_j is a forest (possibly empty), $[\sigma]_i$ denotes the tree obtained by attaching a root with label i to σ, and where the product $[\sigma_1]_{i_1} \cdots [\sigma_\ell]_{i_\ell}$ is unordered (i.e. we consider combinatorial trees). Remark that, $|[\tau]_i| = |\tau| + 1$.

Recall the Butcher elementary differentials $f^\tau \in \mathcal{C}(\mathbb{R}^m, L(\mathbb{R}^n,\mathbb{R}^m))$ of f, indexed by labelled rooted forests τ with $0 \le |\tau| \le N-1$, which are given inductively by $f^{\mathbf{1}} = f$ and then for $\tau = [\sigma_1]_{i_1} \cdots [\sigma_\ell]_{i_\ell}$ by

$$f^\tau = (D^\ell f)(f^{\sigma_1}_{i_1}, \dots, f^{\sigma_\ell}_{i_\ell}) \,, \tag{5.2}$$

where $f^\sigma_i \colon \mathbb{R}^m \to \mathbb{R}^m$ is the i-th component of f^σ for $i \in [n]$, and

$$D^\ell f \colon \mathbb{R}^m \to L((\mathbb{R}^m)^{\otimes \ell}, L(\mathbb{R}^n,\mathbb{R}^m))$$

is the ℓ-th Fréchet derivative of f. Remark that f^τ is a $(|\tau|+1)$-linear function of f involving, in total, $|\tau|$ derivatives of f.

We use the notion of solution to (5.1) from [HK15, Prop. 3.8] (which is the branched analogue of Davie's solutions to geometric RDEs [Dav08, FV08, FV10]). We say that φ solves (5.1) if

$$\varphi_t - \varphi_s = -\varphi_s^p (t-s) + \sum_{0\le|\tau|\le N-1} f^\tau(\varphi_s)\mathbf{X}^{[\tau]}_{s,t} + o(|t-s|) \,, \tag{5.3}$$

where the sum is over all labelled rooted forests τ (possibly empty) with number of vertices $0 \le |\tau| \le N-1$, and where we treat $\mathbf{X}^{[\tau]}$ as $\mathbb{R}^n$-valued with i-th component given by $\mathbf{X}^{[\tau]_i}$.

We make the following assumption for the rest of this section.

Assumption 5.3. For all $0 \le |\tau| < N$ and $0 \le \ell \le N - |\tau|$, there exists $\eta(\tau,\ell) \in \mathbb{R}$ such that

$$\|D^\ell f^\tau\|_{\infty;\eta(\tau,\ell)} < \infty \,,$$

where we recall $\|\cdot\|_{\infty;\eta}$ from (4.15).

As in Section 4.2, we denote in this subsection

$$\alpha = 1/(p-1) \,.$$

Define the local norms, for $z \in (-1,0]$ and $0 < \lambda \le |z+1|$,

$$[\![(f,\mathbf{X})]\!]^{(1)}_{z,\lambda} = \max_{0\le|\tau|\le N-1} H_\tau^{\varrho_\tau} \,, \tag{5.4}$$

where

$$H_\tau = \|f^\tau\|_{\infty;\eta(\tau,0)} \|\mathbf{X}^{[\tau]}\|_{[z-\lambda,z]} \,, \qquad \varrho_\tau = \alpha/\Delta(\tau,0) \tag{5.5}$$

and, for $0 \le \ell \le N - |\tau|$,

$$\Delta(\tau,\ell) = \gamma(|\tau|+1) + \alpha(1-\ell-\eta(\tau,\ell)) \,. \tag{5.6}$$

Define moreover

$$[\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda} = \max_{0\le|\tau|\le N-1} \max_{1\le\ell\le N-|\tau|} \max_{0\le|\sigma|\le N-1} \Big(\|D^\ell f^\tau\|_{\infty;\eta(\tau,\ell)} \|\mathbf{X}^{[\tau]}\|_{[z-\lambda,z]} H_\sigma^{\zeta_{\tau,\ell,\sigma}} \Big)^{\varrho_{\tau,\ell,\sigma}} \tag{5.7}$$

where H_σ is defined in (5.5) and

$$\varrho_{\tau,\ell,\sigma} = \frac{\alpha}{\Delta(\tau,\ell) + \zeta_{\tau,\ell,\sigma}\Delta(\sigma,0)} \,, \qquad \zeta_{\tau,\ell,\sigma} = \frac{(|\tau|+1)\gamma + \ell - 1}{1 - (|\sigma|+1)\gamma} \,,$$

and where, if $\gamma = 1/N$, then the case $|\sigma| = N-1$ is excluded from the max in (5.7) (so, if $\gamma = N = 1$, then simply $[\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda} = 0$). Remark that $\zeta_{\tau,\ell,\sigma} > 0$.

We now make the following assumption for the rest of the section.

Assumption 5.4. ϱ_τ and $\varrho_{\tau,\ell,\sigma}$ are well-defined and positive. Namely, $\Delta(\tau,0) > 0$ and $\Delta(\tau,\ell) + \zeta_{\tau,\ell,\sigma}\Delta(\sigma,0) > 0$ for all $0 \le |\tau| \le N-1$, $0 \le \ell \le N - |\tau|$, and $0 \le |\sigma| \le N-1$ (if $\gamma = 1/N$, then we again consider only $0 \le |\sigma| \le N-2$).

Theorem 5.5. *There exist $C > 0$, depending only on m, n, γ, p and the collection of exponents $\eta(\tau,\ell)$, such that, if $\varphi\colon [-1,0] \to \mathbb{R}^m$ solves (5.1), then, for all $z \in (-1,0]$,*

$$|\varphi_z| \le C\Big([\![(f,\mathbf{X})]\!]^{(1)}_{z,\lambda_z} + [\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda_z} + |z+1|^{-\alpha} \Big) \tag{5.8}$$

where $\lambda_z = (C|\varphi_z|^{-1/\alpha}) \wedge (\frac{1}{2}|z+1|)$.

Remark 5.6. If f is constant, then $D^\ell f^\tau = 0$ unless $\ell = 0$ and $\tau = \mathbf{1}$. Therefore, the right-hand side of (5.8) depends only on the level-1 path $\mathbf{X}^{[\mathbf{1}]}$, as expected from the fact that the higher level components of $\mathbf{X}$ do not affect the solution.

5.2 Simple linear condition on forcing

Before the proof, we derive a consequence of Theorem 5.5 under a *linear* condition on f. Suppose there exist $\eta \in \mathbb{R}$ and $q \geq -1$ such that

$$\|f\|_{\eta,q} \stackrel{\text{def}}{=} \max_{0\leq \ell \leq N} \sup_{x\in\mathbb{R}^m} \frac{|D^\ell f(x)|}{(|x|+1)^{\eta+\ell q}} < \infty \, .$$

A basic example is $f(x) = P(x)\sin(Q(x))$ where P and Q are polynomials of degree η and $q+1$ respectively.

Remark 5.7. The restriction to $q \geq -1$ is natural because, when $q < -1$, we have the equivalence of norms $\|f\|_{\eta,q} \asymp \|f\|_{\eta+Nq+N,-1}$, and therefore we can always reduce to the case $q \geq -1$ modulo changing η.

For $I \subset [-1,0]$, define the homogenous rough path Hölder norm,

$$\|\!|\mathbf{X}|\!\|_I = \max_{1\leq|\tau|\leq N} \|\mathbf{X}^\tau\|_I^{1/|\tau|} \, .$$

Corollary 5.8. *Suppose that*

$$\gamma/\alpha + \gamma - q(1-\gamma) - \eta > 0 \, . \tag{5.9}$$

Then, with notation as in Theorem 5.5, uniformly over all solutions $\varphi\colon [-1,0] \to \mathbb{R}^m$ to (5.1)*,*

$$|\varphi_z| \lesssim \big(\|f\|_{\eta,q}\|\!|\mathbf{X}|\!\|_{[z-\lambda_z,z]}\big)^{\frac{1}{\gamma+\gamma/\alpha-q(1-\gamma)-\eta}} + |z+1|^{-\alpha} \, . \tag{5.10}$$

Remark 5.9. The two cases considered in [BCMW22] (see Assumptions 1.1-1.2 therein) are (i) $q = \eta = 0$, i.e. $f, Df, \dots, D^N f$ are bounded, and (ii) $q = -1$ and $\eta \geq 1$, i.e. f behaves like a polynomial of degree η (see [BCMW22, Rem. 1.9] on the limitation of $\eta \geq 1$). Our condition (5.9) and bound (5.10) agree with [BCMW22, Thm. 1.4] in these two cases, but we make the improvement that we allow any $q, \eta \in \mathbb{R}$ and localise the rough path Hölder norm $\|\!|\mathbf{X}|\!\|$.

Remark 5.10. If $\gamma = 1$, then the above bound does not depend on q, which is consistent with the Young case (4.16) where we have no dependence on the Hölder semi-norm $[f]_\theta$.

Proof. A simple induction shows that

$$\|D^\ell f^\tau\|_{\infty;\eta(\tau,\ell)} \lesssim \|f\|_{\eta,q}^{|\tau|+1} \tag{5.11}$$

where $\eta(\tau,\ell) = (|\tau|+\ell)q + (|\tau|+1)\eta$. Hence Assumption 5.3 is satisfied with this choice of $\eta(\tau,\ell)$ and we have

$$\Delta(\tau,\ell) = \gamma(|\tau|+1) + \alpha(1-\ell-(|\tau|+\ell)q - (|\tau|+1)\eta) \, .$$

Then, by (5.11),

$$[\![(f,\mathbf{X})]\!]_{z,\lambda}^{(1)} \lesssim \max_{0\leq|\tau|\leq N-1} \big(\|f\|_{\eta,q}\|\!|\mathbf{X}|\!\|_{[z-\lambda,z]}\big)^{\frac{\alpha(|\tau|+1)}{\Delta(\tau,0)}} \, . \tag{5.12}$$

Remark that, denoting $k = |\tau|$ with $0 \leq k \leq N-1$,

$$\frac{\Delta(\tau,0)}{\alpha(k+1)} = \frac{\gamma}{\alpha} + \frac{1-kq}{k+1} - \eta \geq \frac{\gamma}{\alpha} + \frac{1}{N} - q + \frac{q}{N} - \eta \geq \frac{\gamma}{\alpha} + \gamma - q(1-\gamma) - \eta \, , \tag{5.13}$$

where we used, in the first equality, that $q \geq -1$ and thus $\frac{1-kq}{k+1}$ is increasing in k and, in the second inequality, that $\gamma \leq 1/N$.

Next, denoting $k = |\tau|, j = |\sigma|$ and recalling $\varrho_{\tau,\ell,\sigma}$ from (5.7), by a quick calculation,

$$\varrho_{\tau,\ell,\sigma} = \frac{1}{\gamma(1+1/\alpha)(k+1+\zeta(j+1)) - \eta(\tau,\ell) - \zeta\eta(\sigma,0)} , \tag{5.14}$$

where we simply denote $\zeta \equiv \zeta_{\tau,\ell,\sigma}$, and

$$\eta(\tau,\ell) + \zeta\eta(\sigma,0) = q(k+\ell+\zeta j) + \eta(k+1+\zeta(j+1)) = (q(1-\gamma)+\eta)(k+1+\zeta(j+1)) .$$

Moreover, due to (5.11),

$$[\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda} \lesssim \max_{\tau,\ell,\sigma} \big(\|f\|_{\eta,q}|\!|\!|\mathbf{X}|\!|\!|_{[z-\lambda,z]}\big)^{(k+1+\zeta(j+1))\varrho_{\tau,\ell,\sigma}} , \tag{5.15}$$

where the exponent, by (5.14), is given by

$$(k+1+\zeta(j+1))\varrho_{\tau,\ell,\sigma} = \frac{1}{\gamma+\gamma/\alpha - q(1-\gamma) - \eta} . \tag{5.16}$$

In conclusion, Assumption 5.4 is equivalent to (5.9), and under this assumption, by (5.12)-(5.13) and (5.15)-(5.16),

$$[\![(f,\mathbf{X})]\!]^{(1)}_{z,\lambda} + [\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda} \lesssim 1 + \big(\|f\|_{\eta,q}|\!|\!|\mathbf{X}|\!|\!|_{[z-\lambda,z]}\big)^{\frac{1}{\gamma+\gamma/\alpha-q(1-\gamma)-\eta}} .$$

Then, by Theorem 5.5, we obtain (5.10). ■

5.3 Proof of Theorem 5.5

We now turn to the proof of Theorem 5.5. The setting is similar to Section 4.2 except that the driver space $\mathbb{M} = \mathbb{M}_{z,\lambda}$ is now all pairs $M = (f,\mathbf{X})$ where $f \in \mathcal{C}^N$ and $\mathbf{X} \in \mathscr{C}^\gamma$ for $\gamma \in (0,1]$, where $\mathscr{C}^\gamma$ is the space of all γ-Hölder branched rough paths.

The rescaling map is defined by

$$R_{z,\lambda}(f,\mathbf{X}) = (f_\lambda, \delta_{\lambda^\alpha}\mathbf{X} \circ T_{z,\lambda})$$

where now $\delta_r\mathbf{X}^\tau = r^{|\tau|}\mathbf{X}^\tau$ and, as before, $f_\lambda = f(\lambda^{-\alpha}\cdot)$.

The solution space $\mathbb{S}_M$ is similarly the set of all solutions to (5.1) in the sense of (5.3). The closure under rescaling $\psi = \lambda^\alpha\varphi \circ T_{z,\lambda}$ assumption (3.4) follows from (5.3) once we note that

$$(f_\lambda)^\tau = \lambda^{-\alpha|\tau|} f^\tau(\lambda^{-\alpha}\cdot)$$

since $|\tau|$ is the total number of derivatives in f^τ, and thus

$$\begin{aligned}\psi_t - \psi_s &= -\lambda^{\alpha-p\alpha}\psi_s^p\lambda(t-s) + \lambda^\alpha \sum_{|\tau|\leq N-1} f^\tau(\lambda^{-\alpha}\psi_s)\mathbf{X}^{[\tau]}_{T_{z,\lambda}(s),T_{z,\lambda}(t)} + o(|t-s|) \\ &= -\psi_s^p(t-s) + \sum_{|\tau|\leq N-1} (f_\lambda)^\tau(\psi_s)(\delta_{\lambda^\alpha}\mathbf{X})^{[\tau]}_{T_{z,\lambda}(s),T_{z,\lambda}(t)} + o(|t-s|) ,\end{aligned}$$

where we used the definition $\alpha = 1/(p-1)$ in the second equality.

It remains to define suitable norms on $\mathbb{M}, \mathbb{M}_{z,\lambda}$ verifying (3.2)-(3.3) and the local coercivity condition (3.5). To this end, we have already introduced the local norms $[\![(f,\mathbf{X})]\!]^{(1)}_{z,\lambda}$ and

$[\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda}$ in (5.4) and (5.7). We then define $[\![(f,\mathbf{X})]\!]_{z,\lambda} = [\![(f,\mathbf{X})]\!]^{(1)}_{z,\lambda} + [\![(f,\mathbf{X})]\!]^{(2)}_{z,\lambda}$ and remark that (3.2) clearly holds for this norm.

Furthermore, we define the norm

$$\|(f,\mathbf{X})\| = \|(f,\mathbf{X})\|_{z,\lambda} = \|(f,\mathbf{X})\|^{(1)} + \|(f,\mathbf{X})\|^{(2)} ,$$

where $\|(f,\mathbf{X})\|^{(i)}$ is defined exactly as $[\![(f,\mathbf{X})]\!]^{(i)}_{z,\lambda}$ but with the Hölder norm $\|\mathbf{X}^\tau\|$ taken over the whole interval $[-1,0]$ (in particular, $\|(f,\mathbf{X})\|^{(i)}$ does not depend on z,λ).

We next verify (3.3). Observe that

$$D^\ell(f_\lambda)^\tau = \lambda^{-\alpha(|\tau|+\ell)}(D^\ell f^\tau)_\lambda$$

and therefore, by Lemma 4.2,

$$\|D^\ell(f_\lambda)^\tau\|_{\infty;\eta(\tau,\ell)} \le \lambda^{-\alpha(|\tau|+\ell+\eta(\tau,\ell))}\|D^\ell f^\tau\|_{\infty;\eta(\tau,\ell)} . \tag{5.17}$$

Furthermore,

$$\|\delta_{\lambda^\alpha}\mathbf{X}^\tau \circ T_{z,\lambda}\|_{|\tau|\gamma} = \lambda^{|\tau|(\alpha+\gamma)}\|\mathbf{X}^\tau\|_{[z-\lambda,z]} .$$

As a consequence

$$\|D^\ell(f_\lambda)^\tau\|_{\infty;\eta(\tau,\ell)}\|\delta_{\lambda^\alpha}\mathbf{X}^{[\tau]}\circ T_{z,\lambda}\|_{(|\tau|+1)\gamma} \le \lambda^{-\alpha(|\tau|+\ell+\eta(\tau,\ell))+(\alpha+\gamma)(|\tau|+1)}\|D^\ell f^\tau\|_{\infty;\eta(\tau,\ell)}\|\mathbf{X}^{[\tau]}\|_{[z-\lambda,z]} .$$

Remark that the exponent of λ is $\Delta(\tau,\ell) = \gamma(|\tau|+1) + \alpha(1-\ell-\eta(\tau,\ell))$ from (5.6). We thus obtain, for $i = 1, 2$,

$$\|R_{z,\lambda}(f,\mathbf{X})\|^{(i)}_{z,\lambda} \le \lambda^\alpha [\![(f,\mathbf{X})]\!]^{(i)}_{z,\lambda}$$

from which (3.3) follows.

To conclude the proof of Theorem 5.5, it remains to verify the form of local coercivity stated in Corollary 3.5. Our argument is self-contained with the exception of the following algebraic and sewing-type lemma. Fix for the rest of the section a solution φ to (5.1).

Lemma 5.11. *For every $0 \le |\tau| \le N-1$, define the 'remainders' $R^\tau_{s,t} \in L(\mathbb{R}^n,\mathbb{R}^m)$ by*

$$f^\tau(\varphi_t) = \sum_{\ell=0}^{N-|\tau|-1} \sum_{(\sigma_1,\dots,\sigma_\ell)} \frac{(D^\ell f^\tau)}{\ell!}(\varphi_s)(f^{\sigma_1}(\varphi_s)\mathbf{X}^{[\sigma_1]}_{s,t},\dots,f^{\sigma_\ell}(\varphi_s)\mathbf{X}^{[\sigma_\ell]}_{s,t}) + R^\tau_{s,t} \tag{5.18}$$

where the inner sum is over all ordered collections of (possibly empty) forests $(\sigma_1,\dots,\sigma_\ell)$ such that $\sum_{i=1}^\ell(|\sigma_i|+1) < N-|\tau|$, and where we recall the notation in (5.2)-(5.3). Then, for all $h \in (0,1]$,

$$\|R^\tau\|_{\gamma(N-|\tau|);\le h} = \sup_{0<|t-s|\le h} \frac{|R^\tau_{s,t}|}{|t-s|^{\gamma(N-|\tau|)}} < \infty . \tag{5.19}$$

Moreover, for $-1 \le s \le t \le 0$, define the rough integral

$$\int_s^t f(\varphi_u)\,\mathrm{d}\mathbf{X}_u = \varphi_t - \varphi_s + \int_s^t \varphi^p_u\,\mathrm{d}u . \tag{5.20}$$

Then we have the 'sewing' estimate for $|t-s| \le h$

$$\Big|\int_s^t f(\varphi_u)\,\mathrm{d}\mathbf{X}_u - \sum_{0\le|\tau|\le N-1} f^\tau(\varphi_s)\mathbf{X}^{[\tau]}_{s,t}\Big| \lesssim |t-s|^{(N+1)\gamma} \max_{0\le|\tau|\le N-1}\|R^\tau\|_{\gamma(N-|\tau|);\le h}\|\mathbf{X}^{[\tau]}\| . \tag{5.21}$$

Proof. The statement is equivalent to saying that $\frac{1}{\ell!}(D^\ell f^\tau)(\varphi_s)(f_{i_1}^{\sigma_1}(\varphi_s),\dots,f_{i_\ell}^{\sigma_\ell}(\varphi_s))$ are the derivatives with respect to $\mathbf{X}$ of the controlled rough path $f(\varphi)$ and $R^\tau_{s,t}$ are the corresponding remainders. For a proof that φ solves (5.1) in the sense of (5.3) if and only if φ is a controlled rough path with derivatives $f^\tau(\varphi)$, see [HK15, Prop. 3.8]. Then, for a proof that $f(\varphi)$ has the claimed derivatives and remainders, see [BCMW22, Cor. 3.9] or the more general statement in the setting of regularity structures in Section 6.6. ∎

Remark 5.12. Since φ solves (5.1) in the sense of (5.3), we have

$$\Big|\int_s^t f(\varphi_u)\,\mathrm{d}\mathbf{X}_u - \sum_{0\le|\tau|\le N-1} f^\tau(\varphi_s)\mathbf{X}^{[\tau]}_{s,t}\Big| = o(|t-s|)$$

from which we see that $\int_s^t f(\varphi_u)\,\mathrm{d}\mathbf{X}_u$ is the usual rough integral in the sense of [Gub04, FdLP06, Gub10, FH20]. The bound (5.21) for $N=1$ is Young's estimate (4.8) (for $\theta=1$ therein). See, for instance, [Gub10, HK15, BCFP19, FH20, BCMW22, Che25] for more background on RDEs.

Remark 5.13. We do not need to be very careful with the regularity of R^τ – our argument works and yields the same a priori bound if we replace $\gamma(N-|\tau|)$ in (5.19) by $\zeta-(|\tau|+1)\gamma+1$ for any $\zeta\in(0,\gamma(N+1)-1]$ and correspondingly replace $(N+1)\gamma$ in (5.21) by $1+\zeta$.

Let us now fix $\varepsilon>0$ and suppose $\|\varphi\|_\infty\le\varepsilon$. By the same argument as above (4.16), it suffices to consider the case that $|\varphi_t|\in[\frac{\varepsilon}{2},\varepsilon]$ for all $t\in[-1,0]$. It suffices therefore to consider the case that $\|D^\ell f\|_\infty<\infty$ for all $0\le\ell\le N$ and $\eta(\tau,\ell)=0$ in Assumption 5.3.

Consider $\mathring{\varphi}\colon[-1,0]\to\mathbb{R}^m$ solving (4.1) with $\mathbf{X}^\tau\equiv 0$ for all non-empty forests $|\tau|>0$ and with the same initial condition $\mathring{\varphi}_{-1}=\varphi_{-1}$. Then $\|\mathring{\varphi}\|_\infty\le\varepsilon$ and there exists $\delta>0$ such that $|\mathring{\varphi}_0|<\varepsilon-2\delta$. Moreover

$$\varphi_t-\mathring{\varphi}_t=\varphi_s-\mathring{\varphi}_s-\int_s^t(\varphi_u^p-\mathring{\varphi}_u^p)\,\mathrm{d}u+\int_s^t f(\varphi_t)\,\mathrm{d}\mathbf{X}_t\ .$$

Since $p>1$ and $\|\varphi\|_\infty,\|\mathring{\varphi}\|_\infty\le\varepsilon$, we have $|\varphi_u^p-\mathring{\varphi}_u^p|\lesssim\varepsilon^{p-1}|\varphi_u-\mathring{\varphi}_u|$ and therefore, provided that $\varepsilon>0$ is sufficiently small,

$$\|\varphi-\mathring{\varphi}\|_\infty\lesssim\Big\|\int_{-1}^{\cdot}f(\varphi_t)\,\mathrm{d}\mathbf{X}_t\Big\|_\infty\ . \tag{5.22}$$

To show local coercivity (3.5), it thus suffices to show that $\|\int_{-1}^{\cdot}f(\varphi_t)\,\mathrm{d}\mathbf{X}_t\|\to 0$ as $\|(f,\mathbf{X})\|\to 0$.

To this end, for all $h\in[0,1]$ by (5.21) and the triangle inequality applied to increments over a mesh of size h as in (4.9), we obtain

$$\Big\|\int_{-1}^{\cdot}f(\varphi_t)\,\mathrm{d}\mathbf{X}_t\Big\|_\infty\lesssim\frac{1}{h}\Big(W_h+V_h\Big) \tag{5.23}$$

where

$$\begin{aligned}W_h&=\max_{0\le|\tau|\le N-1}h^{\gamma(|\tau|+1)}\|f^\tau\|_\infty\|\mathbf{X}^{[\tau]}\|\ ,\\ V_h&=h^{\gamma(N+1)}\max_{0\le|\tau|\le N-1}\|R^\tau\|_{\gamma(N-|\tau|);\le h}\|\mathbf{X}^{[\tau]}\|\ ,\end{aligned}$$

and R^τ is defined by (5.18) in Lemma 5.11. (For $h=0$, we interpret (5.23) as the limit $h\downarrow 0$.)

Note that if $\max_{0\le|\tau|\le N-1}\|f^\tau\|_\infty\|\mathbf{X}^{[\tau]}\|=0$, then $\varphi=\mathring{\varphi}$ by (5.3) and we are done. To exclude this trivial case, we suppose henceforth that

$$\max_{0\le|\tau|\le N-1}\|f^\tau\|_\infty\|\mathbf{X}^{[\tau]}\|>0\ . \tag{5.24}$$

We next estimate $\|R^\tau\|_{\gamma(N-|\tau|);\le h}$ for some fixed $0 \le |\tau| \le N-1$. By Taylor's theorem,

$$f^\tau(\varphi_t) = \sum_{\ell=0}^{N-|\tau|-1} \frac{(D^\ell f^\tau)}{\ell!}(\varphi_s)(\varphi_t - \varphi_s)^\ell + O(\|D^{N-|\tau|} f^\tau\|_\infty |\varphi_t - \varphi_s|^{N-|\tau|}) \,. \tag{5.25}$$

Define the remainder

$$Q_{s,t} = \varphi_t - \varphi_s - \sum_{0\le|\sigma|\le N-1} f^\sigma(\varphi_s)\mathbf{X}_{s,t}^{[\sigma]} \,.$$

Then, by (5.20)-(5.21), for $|t-s| \le h$,

$$\begin{aligned} |Q_{s,t}| &\lesssim |t-s| + |t-s|^{(N+1)\gamma} h^{-(N+1)\gamma} V_h \\ &\le |t-s|^{(N-|\tau|)\gamma} h^{-(N-|\tau|)\gamma}(h + V_h) \,. \end{aligned} \tag{5.26}$$

Note also that

$$|f^\sigma(\varphi_s)\mathbf{X}_{s,t}^{[\sigma]}| \le |t-s|^{(|\sigma|+1)\gamma} h^{-(|\sigma|+1)\gamma} W_h \tag{5.27}$$

and thus

$$|\varphi_t - \varphi_s| \lesssim |t-s| + |t-s|^\gamma h^{-\gamma} W_h + |t-s|^{(N+1)\gamma} h^{-(N+1)\gamma} V_h \,. \tag{5.28}$$

We now expand each factor $\varphi_t - \varphi_s = Q_{s,t} + \sum_{0\le|\sigma|\le N-1} f^\sigma(\varphi_s)\mathbf{X}_{s,t}^{[\sigma]}$ in (5.25) and equate the result with (5.18). We see that the remainder $R_{s,t}^\tau$ is made of a term of order

$$O(\|D^{N-|\tau|} f^\tau\|_\infty |\varphi_t - \varphi_s|^{N-|\tau|}) \tag{5.29}$$

plus a finite sum of terms which are multiples of

$$(D^\ell f^\tau)(\varphi_s) Q_{s,t}^q \prod_{i=1}^{j} f^{\sigma_i}(\varphi_s)\mathbf{X}_{s,t}^{[\sigma_i]} \tag{5.30}$$

where $q + j = \ell$ for $1 \le \ell < N - |\tau|$ and such that

$$q(N-|\tau|) + \sum_{i=1}^{j}(|\sigma_i| + 1) \ge N - |\tau| \,.$$

(The final bound is always fulfilled if $q \ge 1$.) From (5.26)-(5.27), each term (5.30) is bounded by $\|D^\ell f^\tau\|_\infty (h + V_h + W_h)^\ell |t-s|^\theta h^{-\theta}$ with $\theta \ge (N-|\tau|)\gamma$. Combining with (5.28) to control (5.29), we obtain

$$\|R^\tau\|_{(N-|\tau|)\gamma;\le h} \lesssim \max_{1\le\ell\le N-|\tau|} \|D^\ell f^\tau\|_\infty (h + W_h + V_h)^\ell h^{-(N-|\tau|)\gamma} \,.$$

Multiplying by $h^{(N+1)\gamma}\|\mathbf{X}^{[\tau]}\|$ and taking the max over $|\tau| \le N-1$, we obtain

$$V_h \lesssim \max_{0\le|\tau|\le N-1} \max_{1\le\ell\le N-|\tau|} \|D^\ell f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\| h^{(|\tau|+1)\gamma} (h + W_h + V_h)^\ell \,. \tag{5.31}$$

Define

$$Z_h = \max_{0\le|\tau|\le N-1} \max_{1\le\ell\le N-|\tau|} \|D^\ell f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\| h^{(|\tau|+1)\gamma} (h + W_h)^{\ell-1} \,.$$

Lemma 5.14. *For every $\kappa > 0$ there exists $\varepsilon_0 > 0$ such that, if $h \in [0,1]$ and $Z_h < \varepsilon_0$, then*

$$V_h \le \kappa(h + W_h) \,.$$

Proof. It suffices to consider $\kappa \in (0,1]$. By assumption (5.24), $W_h \gtrsim h^{N\gamma}$ and $V_h \lesssim h^{(N+1)\gamma}$, where the proportionality constants can depend on $f, \mathbf{X}, \varphi$. It follows that, $V_h \le \kappa(h + W_h)$ for all h sufficiently small.

On the other hand, if $V_h \le \kappa(h + W_h)$, then by (5.31)

$$V_h \le CZ_h(h + W_h) ,$$

where $C > 0$ does not depend on $f, \mathbf{X}, \varphi$, and therefore, if $Z_h < \varepsilon_0 \stackrel{\text{def}}{=} \kappa C^{-1}/2$, then

$$V_h \le \kappa(h + W_h)/2 .$$

Since $V_h \le \kappa(h + W_h)$ for h sufficiently small and since h, W_h, V_h, Z_h are all continuous and increasing in h, it follows by continuity that $V_h \le \kappa(h + W_h)/2$ for all h such that $Z_h < \varepsilon_0$. ■

Recall that we assume the exponents in Assumption 5.3 are $\eta(\tau, \ell) = 0$. Recall also the definition $\|(f, \mathbf{X})\| = \|(f, \mathbf{X})\|^{(1)} + \|(f, \mathbf{X})\|^{(2)}$ where $\|(f, \mathbf{X})\|^{(1)} = \max_{0\le|\tau|\le N-1}(\|f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\|)^{\varrho_\tau}$ with $\varrho_\tau > 0$ (see (5.5)). It follows immediately that smallness of $\|(f, \mathbf{X})\|^{(1)}$ implies smallness of $\max_{0\le|\tau|\le N-1} \|f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\|$.

Following the approach in the Young case in Section 4.2.2, we now choose the smallest $h \in [0, 1]$ such that $h^{-1}W_h \le \kappa$ for small $\kappa > 0$ (for $h = 0$, we interpret $h^{-1}W_h$ as a limit $h \downarrow 0$); such $h \in [0, 1]$ exists provided $\max_{0\le|\tau|\le N-1} \|f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\| \le \kappa$.

Suppose first $h = 0$. In this case $h^{-1}V_h = 0$ (understood as the limit $h \downarrow 0$), and thus, by (5.23), $\| \int_{-1}^{\cdot} f(\varphi_t)\, \mathrm{d}\mathbf{X}_t\|_\infty \lesssim \kappa$. Hence, if κ is small, then by (5.22), $\|\varphi - \mathring{\varphi}\|_\infty < \delta$, which implies $|\varphi(0)| < \varepsilon - \delta$. This implies local coercivity in the form of Corollary 3.5. Note that this handles $\gamma = 1$ since we always have $h = 0$ in this case.

Suppose now $h \in (0, 1]$. Assume for now that $Z_h < \varepsilon_0$ for ε_0 as in Lemma 5.14. Then $V_h \le \kappa(h + W_h)$ and thus $h^{-1}V_h \le 2\kappa$. Hence, by (5.23), $\| \int_{-1}^{\cdot} f(\varphi_t)\, \mathrm{d}\mathbf{X}_t\|_\infty$ is small and we conclude in the same way as in the case $h = 0$. Therefore, to complete the verification of local coercivity, it remains to show that $Z_h \downarrow 0$ as we take $\|(f, \mathbf{X})\| \downarrow 0$.

Recall the definition

$$\|(f, \mathbf{X})\|^{(2)} = \max_{0\le|\tau|\le N-1} \max_{1\le\ell\le N-|\tau|} \max_{0\le|\sigma|\le N-1} \Big(\|D^\ell f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\| h_\sigma^{(|\tau|+1)\gamma+\ell-1}\Big)^{\varrho_{\tau,\ell,\sigma}}$$

where $\varrho_{\tau,\ell,\sigma} > 0$, $h_\sigma = H_\sigma^{\frac{1}{1-(|\sigma|+1)\gamma}}$, and we recall $H_\sigma = \|f^\sigma\|_\infty \|\mathbf{X}^{[\sigma]}\|$. Furthermore, in the case that $\gamma = \frac{1}{N}$, we exclude $|\sigma| = N - 1$ from the max over σ.

Recall that we took the smallest $h \in [0, 1]$ such that $h^{-1}W_h \le \kappa$ for $\kappa > 0$ small. Since we assume $h > 0$, we must have $h^{-1}W_h = \kappa$ and thus

$$h = (\kappa^{-1}H_\sigma)^{\frac{1}{1-(|\sigma|+1)\gamma}} \lesssim h_\sigma$$

for some $0 \le |\sigma| \le N - 1$; moreover, since $h > 0$, if $\gamma = 1/N < 1$, then this necessarily holds for some $|\sigma| \le N - 2$ (the case $\gamma = N = 1$ is handled above). Consequently, for this $h \in (0, 1]$,

$$Z_h \lesssim \max_{0\le|\tau|\le N-1} \max_{1\le\ell\le N-|\tau|} \|D^\ell f^\tau\|_\infty \|\mathbf{X}^{[\tau]}\| h_\sigma^{(|\tau|+1)\gamma+\ell-1} ,$$

where the right-hand side is small whenever $\|(f, \mathbf{X})\|^{(2)}$ is small. This concludes the verification of local coercivity as stated in Corollary 3.5 and thus also the proof of Theorem 5.5.

6 Singular PDEs

In this section, we derive a priori estimates on coercive PDEs in the framework of regularity structures [Hai14].

6.1 Setup and regularity structures

We use notation from Section 1.1.2.

6.1.1 Functions

For a multi-index $k = (k_1, \dots, k_d) \in \mathbb{N}^d$, denote $|k|_{\mathfrak{s}} = \sum_{i=1}^d k_i \mathfrak{s}_i$. Define

$$\mathbb{N}^d_{<\omega} = \{k \in \mathbb{N}^d \,:\, |k|_{\mathfrak{s}} < \omega\}$$

and analogously for $\mathbb{N}^d_{\leq\omega}, \mathbb{N}^d_{=\omega}$ and so forth.

Let $(E, |\cdot|)$ be a finite-dimensional normed space, which will be the target space of our solution φ. We denote

$$\mathcal{J} = E^{\mathbb{N}^d_{<2}} ,$$

which we think of as the space of jets of order 1 of the solution φ. We write elements of $\mathcal{J}$ in bold font as $\boldsymbol{\varphi} = (\nabla^k \varphi)_{|k|_{\mathfrak{s}}<2}$ or $\boldsymbol{\varphi} = (\varphi, \nabla\varphi)$ with $\nabla\varphi = (\nabla^k\varphi)_{|k|_{\mathfrak{s}}=1} \in E^{\mathbb{N}^d_{=1}}$.

Let $P\colon \mathcal{J} \to E$ be a polynomial which we assume takes the heuristic form

$$P(\boldsymbol{\varphi}) = \varphi^p + B(\nabla\varphi, \varphi^q)$$

where B is bilinear and, if $B \neq 0$, then $2q + 1 = p$. More precisely, we make the following

Assumption 6.1. There exist $p \geq 2$ and a linear map $F\colon E^{\otimes p} \to E$ such that, either $P(\boldsymbol{\varphi}) = F(\varphi^{\otimes p})$, or $q = \frac{p-1}{2}$ is an integer and there exists a linear map $B\colon E^{\mathbb{N}^d_{=1}} \otimes E^{\otimes q} \to E$ such that $P(\boldsymbol{\varphi}) = F(\varphi^{\otimes p}) + B(\nabla\varphi \otimes \varphi^{\otimes q})$. We define

$$\alpha = 2/(p-1) .$$

Remark 6.2. Another way of stating the above assumption on P is that there exists $\alpha > 0$ such that, if we set the degree of $\nabla^k\varphi$ to $-\alpha - |k|_{\mathfrak{s}}$, then P is homogenous of degree $-\alpha - 2$, i.e. the degree of every monomial $\prod_k \nabla^k\varphi$ of P is $\sum_k(-\alpha - |k|_{\mathfrak{s}}) = -\alpha - 2$.

Remark 6.3. Similar to Remark 5.1, it is not crucial for us that P is a polynomial (although it does simplify the argument). But in the case that P is a polynomial, it *is* crucial that the dependence of P on $\nabla\varphi$ is affine. A monomial in P of the form $(\nabla\varphi)^r$ with $r \geq 2$ would require us to take $r(-\alpha - 1) = -\alpha - 2$, thus $\alpha \leq 0$, which is not allowed in our argument.

Assumption 6.1, together with implicit function theorem and standard existence results to parabolic equations, e.g. [Lie96, Thm. 5.6], implies that, for all $\varepsilon \ll 1$ and continuous boundary data $\bar\varphi\colon \partial\Omega \to E$ with $\|\bar\varphi\|_{\infty;\partial\Omega} \leq \varepsilon$, there exists a unique $\mathring\varphi \in \mathcal{C}(\operatorname{cl}\Omega, E)$, smooth on Ω, such that $\mathcal{L}\mathring\varphi = P(\mathring\varphi)$ on Ω and $\mathring\varphi \restriction_{\partial\Omega} = \bar\varphi$. We now make the following additional assumption (which will only be used in Section 6.10) that this solution exhibits coercivity.

Assumption 6.4 (Coercivity)**.** For all $0 < \varepsilon \ll 1$, there exists $\delta > 0$ such that, if $\bar\varphi$ and $\mathring\varphi$ are as above, then $|\mathring\varphi(0)| < \varepsilon - \delta$.

Remark 6.5. We can relax Assumption 6.4 in certain cases. For example, with $E = \mathbb{R}$, the polynomial $P(\boldsymbol{\varphi}) = -\varphi^p$ for even $p \geq 2$ does *not* satisfy Assumption 6.4, but it does satisfy the same coercivity property for *positive* $\bar\varphi$. In this case, a trivial modification of our argument, namely restricting the set of solutions in Section 6.4 to *positive* functions, yields a priori estimates for the equation (such P appeared in [JP25], see Remark 6.42).

Example 6.6. Consider $E = \mathbb{R}^m$ with the Euclidean norm. Then we can take $P(\boldsymbol{\varphi}) = -|\varphi|^{p-1}\varphi$ or $P(\boldsymbol{\varphi}) = (v_1, \dots, v_m)$ where $v_i = -|\varphi_i|^{p-1}\varphi_i$ and $p \geq 3$ is odd.

For an example involving derivatives, consider a linear map $Q\colon E^{\mathbb{N}^d_{=1}} \to \mathfrak{o}(E)$, where $\mathfrak{o}(E)$ is the space of all $A \in L(E,E)$ such that $\langle Ax, x\rangle = 0$ for all $x \in E$, where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product. Set

$$P(\boldsymbol{\varphi}) = -|\varphi|^2\varphi + Q(\nabla\varphi)\varphi \, .$$

Then P satisfies the above assumptions. Indeed, Assumption 6.1 holds with $p = 3$. To verify Assumption 6.4, since $\Delta|\varphi|^2 = 2\langle \Delta\varphi, \varphi\rangle + 2|\nabla\varphi|^2$ and thus $\mathcal{L}|\varphi|^2 = 2\langle \mathcal{L}\varphi, \varphi\rangle - 2|\nabla\varphi|^2$, we have

$$\mathcal{L}|\mathring{\varphi}|^2 = 2\langle -|\mathring{\varphi}|^2\mathring{\varphi} + Q(\nabla\mathring{\varphi})\mathring{\varphi}, \mathring{\varphi}\rangle - 2|\nabla\mathring{\varphi}|^2 = -2|\mathring{\varphi}|^4 - 2|\nabla\mathring{\varphi}|^2 \leq -2|\mathring{\varphi}|^4 \, ,$$

from which the desired condition follows readily from the maximum principle.

6.1.2 Full equation

Recall the maps $T_{z,\lambda}\colon \mathbb{R}^d \to \mathbb{R}^d$ from (1.4). For $a > 0$, denote

$$\Omega_a = T_{0,a}(\Omega) = (-a^2, 0) \times (-a, a)^{d-1} \, , \qquad \partial\Omega_a = T_{0,a}(\partial\Omega) \, .$$

We now consider the equation, posed for $\varphi\colon \Omega_a \to E$,

$$\mathcal{L}\varphi = P(\boldsymbol{\varphi}) + f(\boldsymbol{\varphi})\xi \tag{6.1}$$

where $\boldsymbol{\varphi} = (\nabla^k\varphi)_{|k|_{\mathfrak{s}}\leq 1}\colon \Omega_a \to \mathcal{J}$ is the jet of φ. Here ξ is an $\mathbb{R}^n$-valued distribution in $\mathcal{C}^\beta(\Omega_a)$ for some $\beta < 0$ and $f\colon \mathcal{J} \to L(\mathbb{R}^n, E)$ is a sufficiently regular function. We often write $f = (f_1, \dots, f_n)$ where $f_i\colon \mathcal{J} \to E$ and thus

$$f(\boldsymbol{\varphi})\xi = \sum_{i=1}^n f_i(\boldsymbol{\varphi})\xi_i \, . \tag{6.2}$$

We now make the following assumption throughout this section.

Assumption 6.7 (Subcriticality)**.** The equation (6.1) is subcritical. More specifically,

1. if f is constant, then we require $\beta > -\alpha - 2$;
2. if f depends only on φ, then we require $\beta > -2$;
3. if f depends on $\nabla\varphi$, then we require $\beta > -1$.

We will later make further growth assumptions on f, see Assumption 6.22.

6.1.3 Trees, elementary differentials, and models

As written, the equation (6.1) is, in general, classically ill-posed unless β is not too negative, e.g. for constant $f = 1$ and $P(\boldsymbol{\varphi}) = -|\varphi|^{p-1}\varphi$, classical well-posedness requires $\beta > -2$, which is much more restrictive than our requirement $\beta > -2 - \alpha = -2 - \frac{2}{p-1}$.

The theory of regularity structures [Hai14] allows one to give meaning to solutions to (6.1) for all values of β indicated above provided that ξ is enhanced to a so-called *model* over a *regularity structure* determined by the equation (6.1). We work directly at the level of models, so the distribution ξ will not appear again beyond the equation (6.1). We do not attempt to fully review the theory of regularity structures, but we review the important definitions and results. See [Hai15, FH20, BH20, Che22] for overviews of the theory.

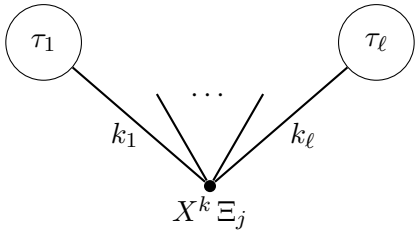


Figure 3: Example of tree notation (6.3)

Definition 6.8 (Trees)**.** A *labelled tree*, or simply *tree*, is a connected, acyclic, undirected graph τ with the following data:

- τ has a distinguished vertex ϱ_τ called the root,
- every vertex v of τ carries a so-called *polynomial* label $\mathfrak{n}(v) \in \mathbb{N}^d$ and a *noise* label $\Xi(v) \in \{0, 1, \dots, n\}$ (recall that n is the dimension of the target space of ξ),
- every edge e carries a *derivative* label $\mathfrak{e}(e) \in \mathbb{N}^d$.

We can write every tree τ uniquely as

$$\tau = X^k \Xi_j \prod_{i=1}^{\ell} \mathcal{I}^{k_i}[\tau_i] \tag{6.3}$$

where the root ϱ_τ has polynomial label $\mathfrak{n}(\varrho_\tau) = k \in \mathbb{N}^d$, noise label $\Xi(\varrho_\tau) = j$, and the factor $\prod_{i=1}^{\ell} \mathcal{I}^{k_i}[\tau_i]$ means that ϱ_τ is incident with $\ell \geq 0$ edges with derivative labels $k_1, \dots, k_\ell$ and where the i-th edge is also incident with the root of another tree τ_i. See Figure 3. Whenever $k_i = 0$, we simply write $\mathcal{I} = \mathcal{I}^{k_i}$. We sometimes write $\Xi_0 = \mathbf{1}$ and drop it from the expression (6.3), suggestive of the fact that Ξ_0 represents the constant '1' function in front of P.

Remark 6.9. Our trees are combinatorial, meaning that the order of edges incident to ϱ_τ is not important, i.e. the product in (6.3) is commutative.

We define the tree factorial $\tau! \in \mathbb{N}$, $\tau! \geq 1$, inductively as follows. Write

$$\tau = X^k \Xi_j \prod_{i=1}^{r} (\mathcal{I}^{k_i}[\tau_i])^{m_i}$$

where $m_i \geq 0$ and $(k_1, \tau_1), \dots, (k_r, \tau_r)$ are pairwise distinct. We then define

$$\tau! = k! \prod_{i=1}^{r} m_i! (\tau_i!)^{m_i} \tag{6.4}$$

where we recall $k! = k_1! \cdots k_d!$ for $k \in \mathbb{N}^d$ and an empty product is equal to 1. We equip the $\mathbb{R}$-linear span of trees with the inner product

$$\langle \tau, \sigma \rangle = \delta_{\tau,\sigma} \tau! \,. \tag{6.5}$$

Definition 6.10. For a tree τ, we define the *homogeneity* $|\tau|_{\mathfrak{s}} \in \mathbb{R}$ as follows. Set

$$|\Xi_j|_{\mathfrak{s}} = \beta \quad \text{for } j = 1, \dots, n \,, \qquad |\mathbf{1}|_{\mathfrak{s}} = 0 \,,$$

and then inductively

$$|X^k\tau|_{\mathfrak{s}} = |k|_{\mathfrak{s}} + |\tau|_{\mathfrak{s}} \,, \quad |\tau\sigma|_{\mathfrak{s}} = |\tau|_{\mathfrak{s}} + |\sigma|_{\mathfrak{s}} \,, \quad |\mathcal{I}^k(\tau)|_{\mathfrak{s}} = |\tau|_{\mathfrak{s}} + 2 - |k|_{\mathfrak{s}} \,.$$

Define also the *critical homogeneity* $\mathfrak{c}(\tau)$ in the same way except $\mathfrak{c}(\Xi_j) = -\alpha - 2 < \beta$.

We next follow the setup of [Hai14, BHZ19, BCCH21] specialised to our equation (6.1) and restrict our space of trees as follows. Denote

$$\widehat{\mathcal{J}} = E^{\mathbb{N}^d} \,, \tag{6.6}$$

which we think of as the space of jets of all order of φ. For a finite dimensional vector space F and $k \geq 0$, let $\mathcal{C}^k(\widehat{\mathcal{J}}, F)$ denote the space of maps $g\colon \widehat{\mathcal{J}} \to F$ that depend on only finitely many components in $E^{\mathbb{N}^d}$ and are of class $\mathcal{C}^k$. Recalling that $\mathcal{J} = \mathbb{N}^d_{<2}$, note that any map $g \in \mathcal{C}^k(\mathcal{J}, F)$ is well-defined as an element of $\mathcal{C}^k(\widehat{\mathcal{J}}, F)$ by projecting to levels $|\ell|_{\mathfrak{s}} < 2$.

Define for $\ell \in \mathbb{N}^d$ and $g \in \mathcal{C}^k(\widehat{\mathcal{J}}, F)$ the function $D^\ell g \in \mathcal{C}^{k-1}(\widehat{\mathcal{J}}, L(E, F))$ as the derivative in the ℓ-th variable, i.e. $D^\ell g(\boldsymbol{\varphi})(h) = D(\boldsymbol{\varphi})(h_\ell)$ where $Dg\colon \widehat{\mathcal{J}} \to L(\widehat{\mathcal{J}}, F)$ is the usual Fréchet derivative and $h_\ell \in E^{\mathbb{N}^d}$ maps ℓ to h and k to 0 for $k \neq \ell$. For $\mathbf{k} = \{k_1, \dots, k_{|\mathbf{k}|}\} \in \mathbb{N}^{\mathbb{N}^d}$ we set

$$D^{\mathbf{k}} g = D^{k_1} \cdots D^{k_{|\mathbf{k}|}} g\colon \mathcal{J} \to L(E^{\otimes |\mathbf{k}|}, F)$$

whenever g is sufficiently regular. (Note that D^ℓ, D^k commute for $\ell, k \in \mathbb{N}^d$, so that $D^{\mathbf{k}}$ is well-defined and compatible with our convention of treating $\mathbf{k}$ as a multi-set.)

For $i \in [d]$ and $g \in \mathcal{C}^k(\widehat{\mathcal{J}}, F)$ define $\partial^i g \in \mathcal{C}^{k-1}(\widehat{\mathcal{J}}, F)$ by

$$\partial^i g(\boldsymbol{\varphi}) = Dg(\boldsymbol{\varphi})(\boldsymbol{\varphi}_{\cdot + i}) = \sum_{\ell \in \mathbb{N}^d} (D^\ell g)(\boldsymbol{\varphi})(\boldsymbol{\varphi}_{\ell + i}) \,,$$

where we use the shorthand i for the singleton $\{i\} \in \mathbb{N}^d$. We then define $\partial^k g$ for any $k \in \mathbb{N}^d$ by composition. (∂^k is well-defined since ∂^i, ∂^j commute. But note that ∂^k, D^ℓ do *not* commute.)

We next define subsets $\mathfrak{R}_j \subset \mathbb{N}^{\mathbb{N}^d}$ for $j \in \{0, 1, \dots, n\}$ (analogous to a 'rule' in [BHZ19], see also [Che22, Sec. 1.3]), as follows. We set

$$\mathfrak{R}_0 = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d} \,:\, D^{\mathbf{k}} P \neq 0\} \,. \tag{6.7}$$

Next, recalling Cases 1-3 in Assumptions 6.7, we define $\mathfrak{R}_j$ for $j > 0$ as follows.

1. If f is constant, then $\mathfrak{R}_j = \emptyset$.
2. If f depends only on φ, then $\mathfrak{R}_j = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d} \,:\, \mathbf{k} \text{ contains only } 0\}$.
3. If f depends on $\nabla\varphi$, then $\mathfrak{R}_j = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d} \,:\, |k|_{\mathfrak{s}} \leq 1 \text{ for all } k \in \mathbf{k}\}$.

Definition 6.11. Consider a tree τ of the form (6.3). We say that τ is *conforming* if each τ_i is conforming and $\{k_1, \dots, k_\ell\} \in \mathfrak{R}_j$. (The base case is $\tau = X^k \Xi_j$, which is conforming for all k, j.) A *leaf* of τ is a non-root vertex $v \neq \varrho_\tau$ such that v is incident with only one edge. We call a leaf v of τ *polynomial* if it has noise label $\Xi(v) = 0$.

- To describe the right-hand side of the equation (6.1), we define the set $\mathfrak{W}$ consisting of all conforming trees τ that have no polynomial leaves. Denote also

  $$\mathfrak{W}_{\leq\omega} = \{\tau \in \mathfrak{W} \,:\, -2 < |\tau|_{\mathfrak{s}} \leq \omega\}$$

  and similarly for $\mathfrak{W}_{<\omega}$. Let $\mathcal{W}_{\leq\omega}$ denote the $\mathbb{R}$-linear span of $\mathfrak{W}_{\leq\omega}$, and similarly for $\mathcal{W}_{<\omega}$.

- To describe the solution φ and its derivatives, we define the set $\mathfrak{U}$ consisting of all trees τ such that $|\tau|_{\mathfrak{s}} > 0$ and $\tau = \mathcal{I}^k[\sigma]$ for some $k \in \mathbb{N}^d$ and $\sigma \in \mathfrak{W}_{<0}$. Remark that $|\tau|_{\mathfrak{s}} < 2$ for every $\tau \in \mathfrak{U}$.

  Consider further the set of polynomials $\bar{\mathfrak{T}} = \{X^k : k \in \mathbb{N}^d\}$. Let $\bar{\mathfrak{T}}_{\leq\omega} = \{X^k : |k|_{\mathfrak{s}} \leq \omega\}$ and denote by $\bar{\mathcal{T}}_{\leq\omega}$ the span of $\bar{\mathfrak{T}}_{\leq\omega}$, and likewise for $\bar{\mathcal{T}}_{<\omega}$. Define the linear space

$$\mathcal{U}_{\leq 2} = \operatorname{span}(\mathfrak{U}) \oplus \bar{\mathcal{T}}_{\leq 2} .$$

- Finally, to put all the relevant trees together, we define the set of trees

$$\mathfrak{T}_{\leq 2} = \mathfrak{W}_{<0} \sqcup \mathfrak{U} \sqcup \bar{\mathfrak{T}}_{\leq 2}$$

  and the linear space

$$\mathcal{T}_{\leq 2} = \operatorname{span}(\mathfrak{T}_{\leq 2}) = \mathcal{W}_{<0} \oplus \mathcal{U}_{\leq 2} .$$

All of the above linear spaces of trees are equipped with the inner product (6.5).

One should keep in mind that $\mathcal{W}_{\leq 0}$ is used to describe the right-hand side of the *remainder* equation of (6.1) and this is the reason that the trees $\tau \in \mathfrak{W}$ with $|\tau|_{\mathfrak{s}} \leq -2$ are excluded from $\mathfrak{W}_{\leq\omega}$. Remark that $\mathfrak{W}_{\leq 0} \cap \bar{\mathfrak{T}}_{\leq 2} = \{\mathbf{1}\}$.

Remark 6.12. Not allowing polynomial leaves in $\mathfrak{W}$ makes our regularity structure (see Definition 6.29) somewhat simpler. The same condition is used in [Hai14], but one could work without it as done in [BHZ19, BCCH21]. The trees in $\mathfrak{U}$ will play a role in the structure group of our regularity structure.

Definition 6.13 (Elementary differentials)**.** For conforming τ and f sufficiently regular, define

$$\Upsilon^\tau \colon \widehat{\mathcal{J}} \to E$$

inductively as follows. Recalling (6.2), we denote for $j = 1, \dots, n$

$$\Upsilon^{\Xi_j} = f_j , \qquad \Upsilon^{\mathbf{1}} = \Upsilon^{\Xi_0} = P .$$

Then for $\tau = X^k \Xi_j \prod_{i=1}^{\ell} \mathcal{I}^{k_i}[\sigma_i]$ we define

$$\Upsilon^\tau(\boldsymbol{\varphi}) = \{\partial^k (D^{k_1} \cdots D^{k_\ell} \Upsilon^{\Xi_j})\}(\boldsymbol{\varphi})(\Upsilon^{\sigma_1}(\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}(\boldsymbol{\varphi})) . \tag{6.8}$$

Remark 6.14. To check that (6.8) make senses, note that $D^{k_1} \cdots D^{k_\ell} \Upsilon^{\Xi_j} \colon \widehat{\mathcal{J}} \to L(E^{\otimes \ell}, E)$, therefore $\{\partial^k (D^{k_1} \cdots D^{k_\ell} \Upsilon^{\Xi_j})\}(\boldsymbol{\varphi}) \in L(E^{\otimes \ell}, E)$, which we evaluate on $(\Upsilon^{\sigma_1}(\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}(\boldsymbol{\varphi}))$ treated as an element of $E^{\otimes \ell}$. Since the derivatives D^{k_i}, D^{k_j} commute, the order of terms in the ℓ-tensor $(\Upsilon^{\sigma_1}(\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}(\boldsymbol{\varphi}))$ does not matter, in agreement with our convention that the product in (6.3) is unordered.

Remark 6.15. If τ is not conforming, then $\Upsilon^\tau = 0$ Moreover, if τ is conforming and Υ^τ depends on $\nabla^\ell \varphi$, then $|\tau|_{\mathfrak{s}} > |\ell|_{\mathfrak{s}} - 2$ by Lemma 6.26 below. In particular, Υ^τ is a well-defined function $\Upsilon^\tau \colon \mathcal{J} \to E$ for all conforming τ with $|\tau|_{\mathfrak{s}} \leq 0$, and Υ^τ is constant for all conforming τ with $|\tau|_{\mathfrak{s}} \leq -2$.

Throughout this section, similar to [Hai14, Assumption 5.3], we make the following standard

Assumption 6.16. Consider $\tau \in \mathfrak{W}_{\leq 0}$ such that $\tau \neq \mathbf{1}$. Then $|\tau|_{\mathfrak{s}}$ is not an integer.

Example 6.17. Suppose β is irrational. Then Assumption 6.16 is satisfied. Indeed, $|\tau|_{\mathfrak{s}} \notin \mathbb{Z}$ for any tree τ with at least one noise label $j \neq 0$. On the other hand, if $\tau \in \mathfrak{W}$ with all noise labels $\Xi(v) = 0$, then, due to Definition 6.11, $\tau = X^k$ for some $k \in \mathbb{N}^d$, which proves the claim.

The next lemma is an important consequence of subcriticality (Assumption 6.7).

Lemma 6.18. *The set $\{\tau : \tau$ is conforming and $|\tau|_{\mathfrak{s}} < \omega\}$ is finite for any $\omega \in \mathbb{R}$. Moreover, if τ is conforming, then $|\tau|_{\mathfrak{s}} \geq \beta$ with equality if and only if $\tau = \Xi_j$ for some $j \neq 0$.*

Proof. This follows from the proof of [Hai14, Lem. 8.10] (see also [BHZ19, Prop. 5.15]). ∎

6.1.4 Growth assumptions on non-linearity

As in the rough path setting, we need to make an assumption on the growth of f and its derivatives. However, while the assumptions on f in Section 5 came only from scaling, here we additionally need to use these assumptions to deal with weighted Schauder estimates. Like in Section 5.1, we first make assumptions directly on the elementary differentials Υ. In Section 6.3 we consider simple linear conditions on f when it behaves like a polynomial.

For $k \in \mathbb{N}^d_{<2}$, let

$$\gamma_k = 1 \wedge \inf\{|\mathcal{I}^k[\tau]|_{\mathfrak{s}} : \mathcal{I}^k[\tau] \in \mathfrak{U}\} \in (0, 1] . \tag{6.9}$$

Remark 6.19. γ_k is a lower bound on the optimal Hölder exponent of $\nabla^k \varphi$, so γ_0 corresponds to γ from Section 5. One has $\gamma_0 = \beta + 2$ if $\beta \in (-2, -1)$ and $\gamma_0 = 1$ if $\beta \in (-1, 0)$. In all cases, $\gamma_k = \gamma_\ell$ if $|k|_{\mathfrak{s}} = |\ell|_s$.

Define further

$$\zeta = \min\{|\tau|_{\mathfrak{s}} - \lfloor |\tau|_{\mathfrak{s}} \rfloor : \tau \in \mathfrak{W} \text{ such that } \lfloor |\tau|_{\mathfrak{s}} \rfloor \leq 0 \text{ and } \tau \neq \mathbf{1}\} . \tag{6.10}$$

Note that $\zeta \in (0, 1)$ by Assumption 6.16 and Lemma 6.18.

Definition 6.20. For a multi-index $\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$, denote $\gamma_{\mathbf{k}} = \sum_{k \in \mathbf{k}} \gamma_k$. For $\tau \in \mathfrak{W}_{\leq 0}$, define

$$\mathring{B}^\tau = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}} : \gamma_{\mathbf{k}} < \zeta - |\tau|_{\mathfrak{s}} \text{ and } D^{\mathbf{k}} \Upsilon^\tau \neq 0\} ,$$
$$B^\tau = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}} : [\mathbf{k} = 0 \text{ or } \mathbf{k} \setminus \{k\} \in \mathring{B}^\tau \text{ for all } k \in \mathbf{k}] \text{ and } D^{\mathbf{k}} \Upsilon^\tau \neq 0\} .$$

Remark 6.21. Our choice for ζ corresponds to $(N+1)\gamma - 1$ in the rough paths case of Section 5.3 (see Remark 6.33) and is, in general, the 'largest possible' in the following sense: if we replace ζ by $\bar{\zeta} \in (0, \zeta]$, then, by considerations similar to the proof of Lemma 6.68, the set of 'solutions' to the PDE (6.1) in the sense of Definition 6.32 does not change.

Recall that we write $\boldsymbol{\varphi} = (\varphi, \nabla\varphi) \in E \oplus E^{\mathbb{N}^d_{=1}}$. For $\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d}$, denote

$$|\mathbf{k}|_{\mathfrak{s}} = \sum_{k \in \mathbf{k}} |k|_{\mathfrak{s}} .$$

We make the following assumption (which will only be used in Sections 6.9-6.10). Recall the notation $|x|_\eta$ from (4.14).

Assumption 6.22. f is sufficiently regular such that $D^{\mathbf{k}} \Upsilon^\tau$ is a continuous function for all $\tau \in \mathfrak{W}_{<0}$ and $\mathbf{k} \in B^\tau$. Moreover, uniformly in $\boldsymbol{\varphi} = (\varphi, \nabla\varphi) \in \mathcal{J}$,

$$|D^{\mathbf{k}} \Upsilon^\tau(\boldsymbol{\varphi})| \lesssim |\varphi|_{\eta(\tau,\mathbf{k})} + |\nabla\varphi|_{\nabla\eta(\tau,\mathbf{k})} |\varphi|_{\bar{\eta}(\tau,\mathbf{k})} , \tag{6.11}$$

where $\eta(\tau,\mathbf{k}), \nabla\eta(\tau,\mathbf{k}), \bar\eta(\tau,\mathbf{k}) \geq 0$ satisfy

$$\Delta(\tau,\mathbf{k}) \stackrel{\text{def}}{=} \alpha - \alpha|\mathbf{k}| + |\tau|_{\mathfrak{s}} + 2 - |\mathbf{k}|_{\mathfrak{s}} - \alpha\eta(\tau,\mathbf{k}) > 0 \,, \tag{6.12}$$

$$\alpha\eta(\tau,\mathbf{k}) = (\alpha+1)\nabla\eta(\tau,\mathbf{k}) + \alpha\bar\eta(\tau,\mathbf{k}) \,, \tag{6.13}$$

$$|\tau|_{\mathfrak{s}} + 2 - |\mathbf{k}|_{\mathfrak{s}} > \nabla\eta(\tau,\mathbf{k}) \,. \tag{6.14}$$

We let $\|\!| D^{\mathbf{k}}\Upsilon^\tau |\!\|$ denote the optimal implicit constant in (6.11) (which depends on $\eta(\tau,\mathbf{k})$ etc.).

A few remarks are in order. The inequality (6.12) is needed to ensure that our 'drivers' shrink under scaling maps $R_{z,\lambda}$ introduced in (6.39) below. The identity (6.13) implies that the two terms on the right-hand side of (6.11) scale in the same way, see Lemma 6.50. The final condition (6.14) is required to close a weighted Schauder estimate in the proof of Lemma 6.68.

Remark 6.23 (PDE vs. ODE)**.** When translating the notation of this section to rough paths (see Remark 6.33), a forest τ in Section 5 corresponds here to a tree $\tilde\tau$ with $|\tilde\tau|_{\mathfrak{s}} = |\tau|\gamma + \gamma - 1$ and $\eta(\tau,\mathbf{k})$ here plays the role of $\eta(\tau,\ell)$ from Assumption 5.3. Since we only consider $|\mathbf{k}|_{\mathfrak{s}} = 0$ in the rough path case, assumption (6.14) is automatically satisfied. Assumption (6.12) becomes equivalent to $\Delta(\tau,\ell) > 0$ in the notation of (5.6). This assumption is simpler but more restrictive than Assumption 5.4 (except when f behaves like a polynomial, in the sense of Section 5.2 with $q = -1$, in which case (6.12) is equivalent to Assumption 5.4). A reason for this gap is that we handle Schauder estimates by integration against the (truncated) heat kernel, which is a highly non-local operation and we could not find a way to extend the localisation argument in Sections 4.2.2 and 5 to the setting here. Furthermore, in contrast to Assumption 5.3, we do not allow $\eta(\tau,\mathbf{k}) < 0$, thus we cannot leverage decay of $D^{\mathbf{k}}\Upsilon^\tau$ at infinity (see Remark 6.71). We find it an interesting problem to determine if these gaps between the two settings can be closed.

Example 6.24. Suppose $f(\boldsymbol{\varphi}) = f(\nabla\varphi)$ depends only on $\nabla\varphi$. If $\beta \in (-1/2, 0)$, in which case (6.1) is classically well-posed, then $\mathfrak{W}_{<0} = \{\Xi_j \,:\, 1 \leq j \leq n\}$ and $\mathring{B}^{\Xi_j} = \{0\}$, and it follows that f satisfies Assumption 6.22 provided there exists $1 \leq \theta < \frac{\alpha+2+\beta}{\alpha+1}$ such that for all $k \in \mathbb{N}^d_{=1}$

$$|f(\nabla\varphi)| \lesssim |\nabla\varphi|_\theta \,, \qquad |D^k f(\nabla\varphi)| \lesssim |\nabla\varphi|_{\theta-1} \,.$$

But if $\beta \in (-1, -1/2)$, then $\gamma_k = 1 + \beta < 1/2$ for $|k|_{\mathfrak{s}} = 1$ and therefore $\mathring{B}^{\Xi_j}$ can contain every singleton $\{k\}$ with $|k|_{\mathfrak{s}} = 1$. It follows that Assumption 6.22 can only be satisfied if $D^{\mathbf{k}} f = 0$ for all $\mathbf{k} = \{k_1, k_2\}$ for which $|\mathbf{k}| = |\mathbf{k}|_{\mathfrak{s}} = 2$. Therefore, in the case $\beta \in (-1, -1/2)$, f must have affine dependence on $\nabla\varphi$.

The next lemma implies that, if f is constant, then Assumption 6.22 is satisfied.

Lemma 6.25. *Consider conforming τ and $\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$. Suppose either f is constant or $\tau = \mathbf{1}$ and define $\eta(\tau,\mathbf{k}), \bar\eta(\tau,\mathbf{k}) \in \mathbb{R}$ by*

$$\alpha\eta(\tau,\mathbf{k}) = 2 + \alpha + \mathfrak{c}(\tau) - |\mathbf{k}|\alpha - |\mathbf{k}|_{\mathfrak{s}} \,, \qquad \alpha + 1 + \alpha\bar\eta(\tau,\mathbf{k}) = \alpha\eta(\tau,\mathbf{k}) \,. \tag{6.15}$$

1. *If $\eta(\tau,\mathbf{k}) < 0$, then $D^{\mathbf{k}}\Upsilon^\tau = 0$.*
2. *If $\bar\eta(\tau,\mathbf{k}) < 0$, then $D^{\mathbf{k}}\Upsilon^\tau$ does not depend on $\nabla\varphi$.*
3. *In all cases, $D^{\mathbf{k}}\Upsilon^\tau$ has affine dependence on $\nabla\varphi$ and*

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim |\varphi|^{\eta(\tau,\mathbf{k})} + \delta_{0,|\mathbf{k}|_{\mathfrak{s}}}|\varphi|^{\bar\eta(\tau,\mathbf{k})}|\nabla\varphi| \,, \tag{6.16}$$

where the final term is absent if $D^{\mathbf{k}}\Upsilon^\tau$ does not depend on $\nabla\varphi$.

Proof. Since either f is constant or $\tau = \mathbf{1}$ by assumption, Υ^τ is a polynomial and $\Upsilon^\tau_\lambda = \Upsilon^\tau$. Then, by Lemma 6.49, $\Upsilon^\tau \circ Q_\lambda = \lambda^{-2-\alpha-\mathfrak{c}(\tau)}\Upsilon^\tau$, thus Υ^τ is homogenous of degree $2+\alpha+\mathfrak{c}(\tau) \geq 0$, where we count $\nabla^k\varphi$ as having degree $\alpha+|k|_\mathfrak{s}$. It is then clear that, if $\eta(\tau,\mathbf{k}) < 0$, then $D^\mathbf{k}\Upsilon^\tau = 0$, while if $\bar\eta(\tau,\mathbf{k}) < 0$, then $D^\mathbf{k}\Upsilon^\tau$ does not depend on $\nabla\varphi$.

Suppose that Υ^τ does not depend on $\nabla\varphi$. Then $D^\mathbf{k}\Upsilon^\tau = 0$ if $|\mathbf{k}|_\mathfrak{s} > 0$ and otherwise

$$|D^\mathbf{k}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim |\varphi|^{\frac{2+\alpha+\mathfrak{c}(\tau)-|\mathbf{k}|\alpha}{\alpha}} = |\varphi|^{\eta(\tau,\mathbf{k})} ,$$

proving (6.16) in this case. Suppose now that Υ^τ depends on $\nabla\varphi$. Then $D^\mathbf{k}\Upsilon^\tau = 0$ if $|\mathbf{k}|_\mathfrak{s} > 1$ since $2(\alpha+1) > 2+\alpha+\mathfrak{c}(\tau)$. It follows that $D^\mathbf{k}\Upsilon^\tau$ has affine dependence on $\nabla\varphi$ and

$$|D^\mathbf{k}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim |\varphi|^{\frac{2+\alpha+\mathfrak{c}(\tau)-|\mathbf{k}|\alpha-|\mathbf{k}|_\mathfrak{s}}{\alpha}} + \delta_{0,|\mathbf{k}|_\mathfrak{s}}|\varphi|^{\frac{1+\mathfrak{c}(\tau)-|\mathbf{k}|\alpha}{\alpha}}|\nabla\varphi| = |\varphi|^{\eta(\tau,\mathbf{k})} + \delta_{0,|\mathbf{k}|_\mathfrak{s}}|\varphi|^{\bar\eta(\tau,\mathbf{k})}|\nabla\varphi| . \quad \blacksquare$$

The following lemma gives several important cases when (6.14) is satisfied with $\nabla\eta(\tau,\mathbf{k}) = 0$.

Lemma 6.26. *Consider a conforming tree τ and $\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d}$ such that $|\mathbf{k}| \geq 1$ and $D^\mathbf{k}\Upsilon^\tau \neq 0$. Assume that either that $|\mathbf{k}| = 1$ or $\beta \leq -1$. Then*

$$|\tau|_\mathfrak{s} > |\mathbf{k}|_\mathfrak{s} - 2 . \tag{6.17}$$

Proof. The case $|\mathbf{k}| = 1$ follows immediately from Lemma 6.58. Suppose now that $\beta \leq -2$. Then, by Lemma 6.58,

$$|\tau|_\mathfrak{s} > |\mathbf{k}|_\mathfrak{s} - |\mathbf{k}|(2+\beta) + \beta \geq |\mathbf{k}|_\mathfrak{s} - 2$$

where the final bound follows from $|\mathbf{k}| \geq 1$ and $2+\beta \leq 0$. Suppose now that $\beta \in (-2,-1]$. Because the set $\mathfrak{W}$ is independent of the choice of $\beta \in (-2,1]$, it suffices to consider $\beta = \varepsilon - 2$ for $\varepsilon > 0$ small. Then, by Lemma 6.58,

$$|\tau|_\mathfrak{s} > |\mathbf{k}|_\mathfrak{s} - |\mathbf{k}|(2+\beta) + \beta = |\mathbf{k}|_\mathfrak{s} - |\mathbf{k}|\varepsilon + \varepsilon - 2 ,$$

and thus $\lim_{\beta\downarrow-2}|\tau|_\mathfrak{s} \geq |\mathbf{k}|_\mathfrak{s} - 2$. Because $|\tau|_\mathfrak{s}$ is increasing in β, we obtain $|\tau|_\mathfrak{s} > |\mathbf{k}|_\mathfrak{s} - 2$. $\blacksquare$

Remark 6.27. For $\beta > -1$ and $|\mathbf{k}| > 1$, (6.17) may fail. E.g. take $E = \mathbb{R}$, $d = 2$, $n = 1$, and $f(\boldsymbol{\varphi}) = \sin(\nabla\varphi)$ so that $D^\mathbf{k}f \neq 0$ for all $\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$ with $|\mathbf{k}| = |\mathbf{k}|_\mathfrak{s}$. Taking $|\mathbf{k}|_\mathfrak{s}$ large, we can then make the right-hand side of (6.17) arbitrarily positive with $|\tau|_\mathfrak{s} = \beta$ fixed.

6.1.5 Models and notion of solution

Denote by $\mathcal{H}^+$ the free commutative algebra generated by X^i and trees $\mathcal{I}^k[\sigma]$ in $\mathfrak{U}$ (see Definition 6.11). We let $\mathbf{1}^+ \in \mathcal{H}^+$ denote the empty product, which is the unit under multiplication. Recall from [Hai14, Sec. 8.1] that there exists a linear map $\Delta\colon \mathcal{T}_{\leq 2} \to \mathcal{H}^+ \otimes \mathcal{T}_{\leq 2}$ characterised by

$$\Delta\Xi_j = \mathbf{1}^+ \otimes \Xi_j , \qquad \Delta X^i = X^i \otimes \mathbf{1} + \mathbf{1}^+ \otimes X^i , \qquad \Delta(\tau\sigma) = (\Delta\tau)(\Delta\sigma)$$

for $j = 0,\dots,n$ and $i \in [d]$, and

$$\Delta\mathcal{I}^k[\sigma] = (\mathrm{Id} \otimes \mathcal{I}^k)\Delta\sigma + \sum_{|\ell|_\mathfrak{s} < |\mathcal{I}^k[\sigma]|_\mathfrak{s}} \mathcal{I}^{k+\ell}[\sigma] \otimes \frac{X^\ell}{\ell!} . \tag{6.18}$$

(Note that we swap the order of tensors compared to [Hai14, BHZ19].)

Remark 6.28. It is possible that $\tau\sigma \in \mathfrak{T}_{\leq 2}$ while τ, σ are not in $\mathfrak{T}_{\leq 2}$. Therefore, the right-hand side of the inductive step $\Delta(\tau\sigma) = (\Delta\tau)(\Delta\sigma)$ should be understood as one term rather than as separate elements $\Delta\tau, \Delta\sigma \in \mathcal{H}^+ \otimes \mathcal{T}_{\leq 2}$. This ensures that Δ is well-defined because every term in the product $(\Delta\tau)(\Delta\sigma)$, upon applying the induction (6.18), is in $\mathcal{H}^+ \otimes \mathfrak{T}_{\leq 2}$.

Every character $g\colon \mathcal{H}^+ \to \mathbb{R}$ (i.e. a multiplicative linear functional for which $g(\mathbf{1}^+) = 1$) defines an invertible linear map $\Gamma\colon \mathcal{T}_{\leq 2} \to \mathcal{T}_{\leq 2}$ by

$$\Gamma\tau = (g \otimes \mathrm{Id})\Delta\tau \ .$$

Note that g can be recovered from Γ by the identities

$$g(\mathcal{I}^k\sigma) = \langle \Gamma\mathcal{I}^k\sigma, \mathbf{1}\rangle \ , \qquad g(X^i) = \langle \Gamma X^i, \mathbf{1}\rangle \tag{6.19}$$

for every $\mathcal{I}^k\sigma \in \mathfrak{U}$ and $i \in [d]$.

We let $\mathcal{G}$ denote the space of linear maps Γ defined in this way by a character $g\colon \mathcal{H}^+ \to \mathbb{R}$. For all $\Gamma \in \mathcal{G}$, the map $\mathrm{Id} - \Gamma$ is triangular in the sense that, for all $\tau \in \mathfrak{T}_{\leq 2}$,

$$\tau - \Gamma\tau \in \mathrm{span}\{\sigma \in \mathfrak{T}_{\leq 2} \,:\, |\sigma|_{\mathfrak{s}} < |\tau|_{\mathfrak{s}}\} \ .$$

Definition 6.29 (Regularity structure and models)**.** Let $\mathcal{A} = \{|\tau|_{\mathfrak{s}} \,:\, \tau \in \mathfrak{T}_{\leq 2}\} \subset (-2, 2]$. The triplet $(\mathcal{T}_{\leq 2}, \mathcal{G}, \mathcal{A})$ is the *regularity structure* associated to the equation (6.1). Recall that $\mathcal{T}_{\leq 2}$ is equipped with the inner product $\langle \cdot, \cdot \rangle$ as in (6.5), for which $\mathfrak{T}_{\leq 2}$ is an orthogonal basis.

Consider an open subset $\mathfrak{K} \subset \mathbb{R}^d$. Recall the notation $\mathfrak{K}^1$ from (1.9). Let $r = \lfloor 1 - \min \mathcal{A} \rfloor \in \{1, 2\}$. A *model* on $\mathfrak{K}$ is a pair $Z = (\Pi, \Gamma)$ where $\Pi = \{\Pi_x\}_{x\in\mathfrak{K}}$ is a family of maps $\Pi_x \in L(\mathcal{T}_{\leq 2}, \mathcal{D}'(\mathfrak{K}^1))$ and $\Gamma = \{\Gamma_{xy}\}_{x,y\in\mathfrak{K}}$ with $\Gamma_{xy} \in \mathcal{G}$ such that

(a) $\Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}$ and $\Pi_y = \Pi_x\Gamma_{xy}$,

(b) for all $\tau \in \mathfrak{T}_{\leq 2}$ and $\mathcal{I}^k[\sigma] \in \mathfrak{U}$

$$\|\Pi\|_{\tau;\mathfrak{K}} \stackrel{\mathrm{def}}{=} \sup_{x\in\mathfrak{K}} \|\Pi_x\tau\|_{|\tau|_{\mathfrak{s}};x;1;r} < \infty \ , \qquad \|\Gamma\|_{\mathcal{I}^k[\sigma];\mathfrak{K}} \stackrel{\mathrm{def}}{=} \sup_{x\neq y\in\mathfrak{K}} \frac{\langle \Gamma_{xy}\mathcal{I}^k[\sigma], \mathbf{1}\rangle}{|x-y|_{\mathfrak{s}}^{|\mathcal{I}^k[\sigma]|_{\mathfrak{s}}}} < \infty \ , \tag{6.20}$$

where we recall the semi-norm $\|\cdot\|_{\omega;x;h;r}$ from (1.8), and

(c) $(\Pi_x X^k\tau)(y) = (y-x)^k(\Pi_x\tau)(y)$, $\Pi_x\mathbf{1} \equiv 1$, and $\Gamma_{xy}X^k = (X - (x-y))^k$.

Let us fix a kernel $K\colon \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$, supported on $B_0(1) \setminus \{0\}$, that agrees with the heat kernel on $B_0(1/2) \setminus \{0\}$ and satisfies Assumptions 5.1 and 5.4 of [Hai14]. Such a kernel exists by [Hai14, Lem. 5.5]. A model is called *adapted to* K if it satisfies the conditions of [Hai14, Def. 5.9], namely, for all $x \in \mathfrak{K}$ and $\mathcal{I}^\ell\tau \in \mathfrak{T}_{\leq 2}$,

$$\Pi_x\mathcal{I}^\ell\tau = D^\ell(K * \Pi_x\tau - \Pi_x\mathcal{J}(x)\tau) \tag{6.21}$$

where the equality is understood in the sense of distributions on $\mathcal{D}'(\mathfrak{K})$ and where

$$\mathcal{J}(x)\tau = \sum_{|k|_{\mathfrak{s}} < |\tau|_{\mathfrak{s}}+2} \frac{X^k}{k!}(D^kK * \Pi_x\tau)(x) \in \bar{\mathcal{T}}_{\leq 2} \ . \tag{6.22}$$

Our notion of model is essentially the same as in [Hai14], the main difference being that we consider a local definition on the set $\mathfrak{K}$ and $\|\Pi\|_{\tau;\mathfrak{K}}$ depends only on $\Pi_x\tau$ tested against functions supported in the *backwards* parabolic ball around $x \in \mathfrak{K}$.

The linear map $\mathcal{I}\colon \mathcal{W}_{<0} \to \mathcal{U}_{\leq 2}$ sending τ to $\mathcal{I}[\tau]$ is an abstract integration map of order 2 in the sense of [Hai14, Def. 5.7].

Remark 6.30. While we do not assume that Π_x is in $\mathcal{D}'(\mathbb{R}^d)$, $K * \Pi_x\tau$ and $\Pi_x\mathcal{J}(x)\tau$ still makes sense as distributions in $\mathcal{D}'(\mathfrak{K})$ for every $x \in \mathfrak{K}$. Indeed, we have a decomposition

$$K = \sum_{n\geq 0} K_n \tag{6.23}$$

where K_n is supported on $B_0(2^{-n})$ and satisfies the assumptions of [Hai14, Sec. 5]. Then for every $N \geq 1$, $K_{<N} = \sum_{n<N} K_n$ is smooth and supported in $B_0(1)$, hence $K_{<N} * \Pi_x\tau$ is a well-defined smooth function on $\mathfrak{K}$ since we assume $\Pi_x \in \mathcal{D}'(\mathfrak{K}^1)$. On the other hand, let $K_{\geq N} = K - K_{<N}$ and $\mathcal{J}_{\geq N}$ be defined as in (6.22) with K replaced by $K_{\geq N}$. Then, for $\psi \in \mathcal{C}_c^\infty(\mathfrak{K})$, it is straightforward to check that there exists N sufficiently large such that $\langle K_{\geq N} * \Pi_x\tau, \psi\rangle$ and $\langle \Pi_x\mathcal{J}_{\geq N}(x)\tau, \psi\rangle$ are both well-defined by using the knowledge of Π_x as an element of $\mathcal{D}'(\mathfrak{K})$ and the bound $\|\Pi\|_{\tau;\mathfrak{K}} < \infty$.

In the sequel, we write $\langle \tau \otimes Y, \sigma\rangle = \langle \tau, \sigma\rangle Y \in E$ for $\tau, \sigma \in \mathcal{T}_{\leq 2}$ and $Y \in E$.

Definition 6.31. Consider a model $Z = (\Pi, \Gamma)$ on an open set $\mathfrak{K} \subset \mathbb{R}^d$ and $\omega \in \mathbb{R}$. An E-valued *modelled distribution* of order ω on $\mathfrak{K}$ is a function $F\colon \mathfrak{K} \to \mathcal{T}_{\leq 2} \otimes E$ such that $\langle F_x, \sigma\rangle = 0$ for $|\sigma|_{\mathfrak{s}} \geq \omega$ and, uniformly in $x, y \in \mathfrak{K}$, $\tau \in \mathfrak{T}_{\leq 2}$, and $|\tau|_{\mathfrak{s}} < \omega$,

$$|\langle F_x, \tau\rangle| \lesssim 1\,, \qquad |\langle F_x - \Gamma_{xy}F_y, \tau\rangle| \lesssim |x-y|_{\mathfrak{s}}^{\omega-|\tau|_{\mathfrak{s}}}\,. \tag{6.24}$$

We denote by $\mathcal{D}^\omega(\mathfrak{K}, E)$ the space of all such functions.

In the theory of regularity structures, one typically lifts the PDE (6.1) to a so-called 'abstract equation' for modelled distributions, which one solves via a fixed point argument. Since we are interested in a priori estimates rather than local well-posedness, we use the following very general notion of 'solution' inspired by Davie's solution to RDEs (5.3). We note that a similar notion appeared independently in the multi-index approach to regularity structures [BOS25].

We are interested in modelled distributions $\Phi\colon \mathfrak{K} \to \mathcal{U}_{\leq 2} \otimes E$ of the form

$$\Phi_x = \boldsymbol{\varphi}_x + \tilde{\boldsymbol{\varphi}}_x\,, \quad \boldsymbol{\varphi}_x = \varphi_x\mathbf{1} + \sum_{0<|k|_{\mathfrak{s}}\leq 2} \nabla^k\varphi_x \frac{X^k}{k!}\,, \quad \tilde{\boldsymbol{\varphi}}_x = \sum_{|\tau|_{\mathfrak{s}}<0} \Upsilon^\tau(\boldsymbol{\varphi}_x)\frac{\mathcal{I}[\tau]}{\tau!}\,, \tag{6.25}$$

where we recall $\tau!$ from (6.4). Here and below, summation over k means $k \in \mathbb{N}^d$ and summation over τ means $\tau \in \mathfrak{T}_{\leq 2}$ subject to the given constraints. We also use the shorthand $Y\tau = \tau \otimes Y$ for $Y \in E$ and $\tau \in \mathcal{T}_{\leq 2}$. Furthermore, $\nabla^k\varphi\colon \mathfrak{K} \to E$ is a function that is not necessarily a derivative of φ and we identify $\bar{\mathcal{T}}_{<2} \otimes E$ with $\mathcal{J} = E^{\mathbb{N}^d_{<2}}$ via $\nabla^k\varphi\frac{X^k}{k!} \mapsto \nabla^k\varphi$ to makes sense of $\Upsilon^\tau(\boldsymbol{\varphi})$. Recall that $\Upsilon^\tau\colon \mathcal{J} \to E$ is well-defined by Remark 6.15.

Such modelled distributions in [BCCH21] are called *coherent*. It was observed therein that solutions to the 'abstract' version of (6.1) are necessarily coherent.

Note that Φ_x is clearly determined by $\boldsymbol{\varphi}_x$. Moreover, by Remark 6.15, if Υ^τ depends on $\nabla^k\varphi$, then $|\mathcal{I}[\tau]|_{\mathfrak{s}} > |k|_{\mathfrak{s}}$. It follows that, for an open set $\mathfrak{K}$, given $\varphi\colon \mathfrak{K} \to E$ and $\omega > 2$, there exists at most one function $\nabla^k\varphi\colon \mathfrak{K} \to E$ for every $|k|_{\mathfrak{s}} \leq 2$ such that

$$|\varphi_x - \langle \mathbf{1}, \Gamma_{xy}\Phi_y\rangle| \lesssim |x-y|_{\mathfrak{s}}^\omega\,,$$

where we write Φ for the function in (6.25). That is, $\nabla^k\varphi$, whenever it exists, is uniquely determined by φ (the coefficient of $\mathbf{1}$) once the model Z is given. In the sequel, we let $\nabla^k\varphi$ denote these functions and let Φ be defined as in (6.25).

Definition 6.32 (Solution)**.** Consider $0 < a \leq 1$ and a model $Z = (\Pi, \Gamma)$ on Ω_a. We say that $\varphi \in \mathcal{C}(\Omega_a, E)$ solves the remainder of (6.1) for Z if Φ is in $\mathcal{D}^{2+\zeta}(\Omega_a, E)$ and

$$\sup_{x \in \Omega_a} \Big\| \mathcal{L}\varphi - \sum_{\tau \in \mathfrak{W}_{\leq 0}} \Upsilon^\tau(\boldsymbol{\varphi}_x) \frac{1}{\tau!} \Pi_x \tau \Big\|_{\zeta; x; \frac{1}{2}\|\!| x |\!|_a; r} < \infty \,, \tag{6.26}$$

where we recall $\zeta \in (0,1)$ from (6.10), r from Definition 6.29, and the localised semi-norms (1.8), and we denote

$$\|\!| x |\!|_a = \inf_{y \in \partial\Omega_a} |x - y|_{\mathfrak{s}} \,.$$

Readers familiar with regularity structures, will realise that (6.26) is essentially the statement that $\mathcal{L}\varphi$ is the reconstruction of $\sum_{\tau \in \mathfrak{W}_{\leq 0}} \Upsilon^\tau(\boldsymbol{\varphi}_x) \frac{1}{\tau!}\tau$ on Ω_a (see also (6.69) below). As mentioned earlier, we restrict to $|\tau|_{\mathfrak{s}} > -2$ because we are looking at the *remainder* of the equation (6.1), while the 'full solution' is given by $\varphi_x + \sum_{\tau \in \mathfrak{W} \,:\, |\tau|_{\mathfrak{s}} \leq -2} \Upsilon^\tau \frac{1}{\tau!} \Pi_x \tau$. The remainder φ is the same quantity that is bounded in [MW20b, CMW23, GH21]. (For every $|\tau|_{\mathfrak{s}} \leq -2$, one has $\Gamma\tau = \tau$ for all $\Gamma \in \mathcal{G}$ thus $\Pi_x \tau = \Pi_y \tau$, and, according to Remark 6.15, Υ^τ is constant.)

Remark 6.33 (Comparison to rough paths)**.** When comparing to Section 5, $\mathbf{X}_{s,t}^{[\tau_1 \cdots \tau_n]_i}$ should be thought as the integral over $[s,t]$ of $\Pi_s(\Xi_i \mathcal{I}[\tau_1] \cdots \mathcal{I}[\tau_n])$. It is therefore natural to identify $\tau = \tau_1 \cdots \tau_n$ with the vector of trees $\tilde{\tau} = (\Xi_i \mathcal{I}[\tau_1] \cdots \mathcal{I}[\tau_n])_{i=1,\ldots,n}$, e.g. $\mathbf{1}$ from Section 5 becomes $(\Xi_1, \ldots, \Xi_n)$. Then, supposing for simplicity that $n = 1$, because the homogeneity of the noise is $\beta = \gamma - 1$ and $\mathcal{I}$ is now 1-regularising, one has $|\tilde{\tau}|_{\mathfrak{s}} = (|\tau| + 1)(\gamma - 1) + |\tau| = |\tau|\gamma + \gamma - 1$, so the exponent $(N - |\tau|)\gamma$ from (5.19) corresponds to $\zeta - |\tilde{\tau}|_{\mathfrak{s}}$ in (6.24).

Remark 6.34. Similar to Remark 5.13, we could work with any $\bar{\zeta} \in (0, \zeta]$ and obtain the same final result. Our choice of ζ in (6.10) is the largest possible in the sense of Remark 6.21.

6.2 Main result

We can now state the main result of this section. Recall that Assumptions 6.1, 6.4, 6.7, 6.16, 6.22 are in place throughout this section.

Consider a function $f \colon \mathcal{J} \to L(\mathbb{R}^n, E)$ satisfying Assumption 6.22. Consider $0 < a \leq 1$ and an adapted model $Z = (\Pi, \Gamma)$ on Ω_a. For $z \in \Omega_a$ and $0 < \lambda \leq \|\!| z |\!|_a$ define

$$[\![(f, Z)]\!]_{z,\lambda} = \max_{\tau \in \mathfrak{W}_{<0}} \max_{\mathbf{k} \in B^\tau} \Big(\|\!| D^{\mathbf{k}} \Upsilon^\tau |\!| \|\Pi\|_{\tau; B_z(\lambda)} \Big)^{\varrho(\tau, \mathbf{k})} + \max_{\mathcal{I}^k[\sigma] \in \mathfrak{U}} \Big(\|\!| \Upsilon^\sigma |\!| \|\Gamma\|_{\mathcal{I}^k[\sigma]; B_z(\lambda)} \Big)^{\varrho(\sigma, 0)} \tag{6.27}$$

where we recall (6.20), $\|\!| D^{\mathbf{k}} \Upsilon^\tau |\!|$ is taken from Assumption 6.22, and we define

$$\varrho(\tau, \mathbf{k}) = \frac{\alpha}{\Delta(\tau, \mathbf{k})} \tag{6.28}$$

where $\Delta(\tau, \mathbf{k}) > 0$ is defined in (6.12) of Assumption 6.22.

Theorem 6.35. *Let $\varphi \in \mathcal{C}(\operatorname{cl}\Omega_a, E)$ solve the remainder of (6.1) in Ω_a in the sense of Definition 6.32. Then there exists $C > 0$, not depending on a, Z, f, φ, such that, for all $z \in \Omega_a$,*

$$|\varphi(z)| \leq C \max\{ [\![(f, Z)]\!]_{z, \lambda_z}, \|\!| z |\!|_a^{-\alpha} \} \,,$$

where $\lambda_z = (C|\varphi_z|^{-1/\alpha}) \wedge (\frac{1}{2}\|\!| z |\!|_a)$.

Remark 6.36. $[\![(f, Z)]\!]_{z,\lambda_z}$ depends on Γ_{xy} only for $x, y \in B_z(\lambda_z)$ and on Π_x evaluated on non-anticipative test functions centred at $x \in B_z(\lambda_z)$ but at the *global* scale 1. The local dependence on Γ is natural (cf. Section 5 for rough paths) but the global dependence on Π we believe is an artefact of the way we handle Schauder estimates in our proof (Lemma 6.63).

6.3 Examples: linear conditions on f

In this subsection, we verify Assumption 6.22 in the case that f behaves like a polynomial and show that the corresponding exponents in $[\![(f,Z)]\!]_{z,\lambda}$ take a simple form.

For $\tau \in \mathfrak{W}_{<0}$, let $L(\tau)$ denote the number of instances of Ξ_j with $j \in \{1,\dots,n\}$ in τ (i.e. the number of 'noises'). Define the homogenous model norm

$$|\!|\!|Z|\!|\!|_{z,\lambda} = \max_{\tau\in\mathfrak{W}_{<0}} \|\Pi\|_{\tau;B_z(\lambda)}^{1/L(\tau)} + \max_{\mathcal{I}^k[\sigma]\in\mathfrak{U}} \|\Gamma\|_{\mathcal{I}^k[\sigma];B_z(\lambda)}^{1/L(\sigma)} .$$

Consider a tree τ of the form (6.3). For vertices v, w of τ, we write $v \preceq w$ if v is on the unique path from w to the root ϱ_τ and write $v \prec w$ if in addition $v \neq w$. We call a tree σ a *branch* of τ if σ has a root ϱ_σ, which is a vertex of τ, and contains all vertices v of τ such that $\varrho_\sigma \preceq v$; all edges and labels of σ are inherited from τ. In other words, σ is a branch of τ if either $\sigma = \tau$ or if σ is a branch of $\tau_1, \dots,$ or τ_ℓ.

Lemma 6.37. *For any conforming tree τ and a branch σ of τ, one has $|\tau|_{\mathfrak{s}} \geq |\sigma|_{\mathfrak{s}}$ with equality if and only if $\sigma = \tau$.*

Proof. Replacing σ by a single vertex with noise label Ξ_1 keeps the tree conforming, from which, by Lemma 6.18, it follows that $|\tau|_{\mathfrak{s}} - |\sigma|_{\mathfrak{s}} + \beta \geq \beta$ with equality if and only if $\tau = \sigma$. ■

In what follows, we let f, Z, φ, λ_z be as in Theorem 6.35.

6.3.1 f constant

Corollary 6.38. *Suppose that f is constant (which is always the case if $\beta < -2$). Then*

$$|\varphi(z)| \lesssim \max\{|\!|\!|Z|\!|\!|_{z,\lambda_z}^{\frac{\alpha}{\beta+\alpha+2}}, |\!|x|\!|_a^{-\alpha}\} .$$

Proof. Assumption 6.22 is satisfied by Lemma 6.25 with $\eta, \bar\eta$ given by (6.15) (except we take $\bar\eta(\tau,\mathbf{k}) = 0$ if $D^{\mathbf{k}}\Upsilon^\tau$ does not depend on $\nabla\varphi$). Then, for any $|\tau|_{\mathfrak{s}} \in \mathfrak{W}_{<0}$ and $\mathbf{k}$,

$$\Delta(\tau,\mathbf{k}) = |\tau|_{\mathfrak{s}} - \mathfrak{c}(\tau) = L(\tau)(\beta+\alpha+2) .$$

The proof follows from Theorem 6.35. ■

Remark 6.39. In the case that $p = 3$ and P does not depend on $\nabla\varphi$, this estimate essentially agrees with [MW20b, CMW23].

6.3.2 f depends only on φ

Let us write $f(\boldsymbol{\varphi}) = f(\varphi, \nabla\varphi)$ and denote by $D_1 f$ and $D_2 f$ the partial derivatives with respect to φ and $\nabla\varphi$ respectively.

Suppose now that f is non-constant and depends only on φ, i.e. $D_2 f = 0$. Since f is non-constant, we have $\beta \in (-2, 0)$. Let $N = \lfloor 2/(2+\beta) \rfloor$ and suppose that

$$[\![f]\!]_\eta = \max_{l=0,\dots,N} \sup_{\boldsymbol{\varphi}\in\mathcal{J}} \frac{|D_1^l f(\boldsymbol{\varphi})|}{|\varphi|_{\eta-l}} < \infty ,$$

where $\eta \geq 0$ and satisfies

$$\alpha(\eta - 1) < \beta + 2 . \tag{6.29}$$

Moreover, if f is not a polynomial, we assume that $\eta \geq N$.

Corollary 6.40. *Under the above assumptions,*

$$|\varphi(z)| \lesssim \max\{([\![f]\!]_\eta \|Z\|_{z,\lambda_z})^{\frac{\alpha}{\beta+2+\alpha-\alpha\eta}}, \|x\|_a^{-\alpha}\} \,. \tag{6.30}$$

Remark 6.41. Remark that, if we substitute $2+\beta \rightsquigarrow \gamma$, then condition (6.29) matches condition (5.9) with $q=-1$, and the exponent $\frac{\alpha}{\beta+2+\alpha-\alpha\eta}$ in (6.30) matches that in (5.10).

Remark 6.42. [JP25, Thm. 4.1] considers the case above with $p=2$, $f(\boldsymbol{\varphi})=\varphi$, and φ a positive solution[1] to $\mathcal{L}\varphi = -\varphi^2 + \varphi\xi$, where $\xi \in \mathcal{C}^\beta$ with $\beta = -1-\varepsilon$ for $\varepsilon > 0$ small. Taking $\eta = 1$, we recover their result restricted to the unit box.[2]

For the proof of Corollary 6.40, we require an extra lemma. For a tree τ and a vertex v of τ, we call an edge $\{v,w\}$ *outgoing* if $v \prec w$.

Lemma 6.43. *Consider $\tau \in \mathfrak{W}_{<0}$. Then there is at most one vertex v of τ with noise label $\Xi(v)=0$. Moreover, v has precisely one outgoing edge e for which $|\mathfrak{e}(e)|_{\mathfrak{s}} = 1$.*

Proof. Consider a vertex v with $\Xi(v)=0$. Then v can have at most one outgoing edge e with $|\mathfrak{e}(e)|_{\mathfrak{s}}=1$ since P is affine in $\nabla\varphi$ and by the choice of $\mathfrak{R}_0$ in (6.7).

Suppose now that $\mathfrak{e}(e)=0$ for all outgoing edges e of v. Let us write σ_v for the unique branch with root v. Then, since $\beta > -2$, we have $|\sigma_v|_{\mathfrak{s}} > 0$, which contradicts $|\tau|_{\mathfrak{s}} < 0$ by Lemma 6.37. Therefore v has precisely one outgoing edge with $|\mathfrak{e}(e)|_{\mathfrak{s}}=1$.

Now suppose there is another vertex $w \neq v$ with $\Xi(w)=0$. Note that $|\sigma_v|_{\mathfrak{s}} > 1+\beta > -1$. If $w \prec v$, then $|\sigma_w|_{\mathfrak{s}} > 0$ because $|\sigma_v|_{\mathfrak{s}} > -1$, which contradicts $|\tau|_{\mathfrak{s}} < 0$ by Lemma 6.37. Likewise if $v \prec w$. Suppose now that $v \not\prec w$ and $w \not\prec v$. We can replace the branches σ_v, σ_w by single vertices Ξ_1 and the resulting tree will be conforming. Therefore, by Lemma 6.18, $|\tau|_{\mathfrak{s}} - |\sigma_v|_{\mathfrak{s}} - |\sigma_w|_{\mathfrak{s}} + 2\beta > \beta$, which implies

$$|\tau|_{\mathfrak{s}} > |\sigma_v|_{\mathfrak{s}} + |\sigma_w|_{\mathfrak{s}} - \beta > 2 + \beta > 0$$

where we used $|\sigma_v|_{\mathfrak{s}}, |\sigma_w|_{\mathfrak{s}} > 1+\beta$ in the final bound. This again contradicts $|\tau|_{\mathfrak{s}} < 0$. ∎

Proof of Corollary 6.40. Consider a tree $\tau \in \mathfrak{W}_{<0}$. We proceed to verify Assumption 6.22.

We first claim that, if $\mathbf{k} \in \mathring{B}^\tau$, then $\mathbf{k} = \{0,\dots,0\}$ contains only 0 with $|\mathbf{k}| \leq N-1$; in particular, if $\beta \in (-1,0)$, thus $N=1$, then $\mathbf{k} = \emptyset$. Indeed, if $|\mathbf{k}|_{\mathfrak{s}} > 0$, then $D^{\mathbf{k}}\Upsilon^\tau = 0$ by Lemma 6.43 and the assumption that P is affine in $\nabla\varphi$ and that f does not depend on $\nabla\varphi$. It thus suffices to consider the case that $\mathbf{k}$ contains only 0.

Suppose $|\mathbf{k}| \geq N$. If $\beta \in (-1,0)$, then $\tau = \Xi_j$ for $j > 0$, and $\gamma_0 = 1$, $\zeta \leq \beta + 1$, hence

$$\gamma_{\mathbf{k}} \geq 1 \geq \zeta - \beta \,,$$

thus $\mathbf{k} \notin \mathring{B}^\tau$. If $\beta \in (-2,-1)$, then $\gamma_0 = 2+\beta$. Note that N is the smallest integer such that $N(2+\beta)+\beta > 0$. Since there exists $\sigma \in \mathfrak{W}$ with $|\sigma|_{\mathfrak{s}} = N(2+\beta)+\beta \in (0,1)$, it follows that $\zeta \leq N(2+\beta)+\beta$. Therefore

$$\gamma_{\mathbf{k}} \geq N(2+\beta) \geq \zeta - \beta \,,$$

so again $\mathbf{k} \notin \mathring{B}^\tau$. This proves the claim.

[1] Although Theorem 6.35, as written, does not apply in this setting, see Remark 6.5 for a trivial fix.

[2] [JP25, Thm 4.1] states an a priori estimate that is *uniform* in boxes of diameter ≥ 1. Such a bound is useful when the driver is small and the box is large. While our result does not directly imply this, our method to prove local coercivity by comparing to the 'zero-noise' solution can be used to prove such a uniform bound. Since we find this argument a separate task, we prefer to avoid further details.

To verify Assumption 6.22, suppose first that τ has no noise label Ξ_0. Using the shorthand $L = L(\tau)$, we have $|\tau|_{\mathfrak{s}} = (L-1)(\beta+2)+\beta$. Moreover, for $\mathbf{k} \in B^\tau$, by the above claim and definition of $[\![f]\!]_\eta$,

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim [\![f]\!]_\eta^L |\varphi|_{\eta L-(L-1+|\mathbf{k}|)} \tag{6.31}$$

Set $\nabla\eta(\tau,\mathbf{k}) = 0$ and $\eta(\tau,\mathbf{k}) = \eta L-(L-1+|\mathbf{k}|)$. By $D^{\mathbf{k}}\Upsilon^\tau \neq 0$ and $|\mathbf{k}| < N$ and the assumption that $\eta \geq N$ if f is not a polynomial, one has $\eta(\tau,\mathbf{k}) \geq 0$. Then condition (6.12) is equivalent to

$$\Delta(\tau,\mathbf{k}) = \alpha(1-|\mathbf{k}|)+(L-1)(\beta+2)+\beta+2-\alpha(\eta L-(L-1+|\mathbf{k}|)) = L(\beta+2-\alpha\eta+\alpha) > 0\,,$$

which holds by (6.29).

By Lemma 6.43, it remains to consider the case that τ has one vertex v with $\Xi(v) = 0$ and with one outgoing edge for which $|\mathfrak{e}(e)|_{\mathfrak{s}} = 1$. Recall by Assumption 6.1 that this can only happen if p is odd and, for any $|k|_{\mathfrak{s}} = 1$, $D^k P(\boldsymbol{\varphi})$ is a homogenous polynomial in φ of degree $\frac{p-1}{2} = 1/\alpha$. In this case, $|\tau|_{\mathfrak{s}} = L(\beta+2)-1$. Then, for $\mathbf{k} \in B^\tau$,

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim [\![f]\!]_\eta^L |\varphi|_{\eta L-(L-1)+1/\alpha-|\mathbf{k}|}\,. \tag{6.32}$$

Set $\nabla\eta(\tau,\mathbf{k}) = 0$ and $\eta(\tau,\mathbf{k}) = \eta L-(L-1)+1/\alpha-|\mathbf{k}|$. Then again $\eta(\tau,\mathbf{k}) \geq 0$ and condition (6.12) is equivalent to

$$\Delta(\tau,\mathbf{k}) = \alpha(1-|\mathbf{k}|)+L(\beta+2)-1+2-\alpha(\eta L-(L-1)+1/\alpha-|\mathbf{k}|) = L(\beta+2-\alpha\eta+\alpha) > 0\,,$$

which is the same condition as earlier and holds by (6.29). The bounds (6.31)-(6.32) imply that, with the above choices for $\eta(\tau,\mathbf{k}), \nabla\eta(\tau,\mathbf{k})$,

$$\|\!| D^{\mathbf{k}}\Upsilon^\tau |\!\| \lesssim [\![f]\!]_\eta^L\,.$$

Finally, note that condition (6.14) is equivalent to $|\tau|_{\mathfrak{s}}+2 > 0$, which clearly holds due to $|\tau|_{\mathfrak{s}} \geq \beta$ by Lemma 6.18. Therefore Assumption 6.22 holds and Theorem 6.35 yields the result. ∎

Remark 6.44. Suppose we instead have the more general bound, similar to Section 5.2,

$$[\![f]\!]_{\eta,q} = \sup_{l,\boldsymbol{\varphi}} \frac{|D_1^l f(\boldsymbol{\varphi})|}{|\varphi|_{\eta+lq}} < \infty$$

for some $q \in \mathbb{R}$. Then, in the first case in the above proof, instead of (6.31) one would have

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim [\![f]\!]_{\eta,q}^L |\varphi|_{\eta L+(L-1+|\mathbf{k}|)q}\,.$$

It is then natural to set $\nabla\eta(\tau,\mathbf{k}) = 0$ and $\eta(\tau,\mathbf{k}) = \eta L+(L-1+|\mathbf{k}|)q$ for which

$$\Delta(\tau,\mathbf{k}) = L(\beta+2-\alpha\eta-\alpha q)-\alpha|\mathbf{k}|(q+1)\,,$$

and Assumption 6.22 is satisfied if this expression is positive for all $|\mathbf{k}|, L \leq N$. Similar considerations apply to the second case in the proof of Corollary 6.40. However, $\Delta(\tau,\mathbf{k})$ would no longer be a multiple of L for $q \neq -1$, so it appears the best bound that Theorem 6.35 provides in this case, in contrast to Section 5.2, is not simply a power of $[\![f]\!]_{\eta,q}\|\!| Z |\!\|$.

6.3.3 f depends on $\nabla\varphi$

Suppose now that f depends on $\nabla\varphi$, which implies $\beta \in (-1, 0)$. Writing $f(\boldsymbol{\varphi}) = f(\varphi, \nabla\varphi)$, we denote by $D_1 f$ and $D_2 f$ the partial derivatives with respect to φ and $\nabla\varphi$ respectively. By Example 6.24, unless the equation is classically well-posed ($\beta > -\frac{1}{2}$), the dependence of f on $\nabla\varphi$ must be affine for Assumption 6.22 to hold, i.e. $D_2^2 f = 0$, which we now assume.

Denoting $\theta = \frac{\alpha+1}{\alpha}$, suppose that

$$\vertiii{f}_\eta = \max_{l=0,1} \sup_{\boldsymbol{\varphi}\in\mathcal{J}} \frac{|D_1^l f(\boldsymbol{\varphi})|}{|\varphi|_{\eta-l} + |\nabla\varphi|_1 |\varphi|_{\eta-\theta-l}} + \max_{l=0,1} \sup_{\boldsymbol{\varphi}\in\mathcal{J}} \frac{|D_2 D_1^l f(\boldsymbol{\varphi})|}{|\varphi|_{\eta-\theta-l}} < \infty \,,$$

where η satisfies (6.29) and $\eta \geq \theta$. If f is not a polynomial, we further assume that $\eta \geq \theta + 1$.

Corollary 6.45. *Under the above assumptions,*

$$|\varphi(z)| \lesssim \max\{(\vertiii{f}_\eta \vvvert Z \vvvert_{z,\lambda_z})^{\frac{\alpha}{\beta+2+\alpha-\alpha\eta}}, \vvvert x \vvvert_a^{-\alpha}\} \,.$$

Proof. Consider a tree $\tau \in \mathfrak{W}_{<0}$. By Lemma 6.37, all noise labels of τ must be Ξ_j for $j > 0$ and all derivative labels must satisfy $|\mathfrak{e}(e)|_{\mathfrak{s}} = 1$ (a noise label Ξ_0 or edge e with $\mathfrak{e}(e) = 0$ implies the existence of a branch σ with $|\sigma|_{\mathfrak{s}} > 0$ because $\beta > -1$). Since we suppose f is affine in $\nabla\varphi$, we have $\Upsilon^\tau \neq 0$ only if $\tau = \Xi_{j_1}\mathcal{I}^{k_1}[\cdots\mathcal{I}^{k_{L-1}}[\Xi_{j_L}]]$ with $|k_i|_{\mathfrak{s}} = 1$ and $j_i \geq 1$ and where we denote $L = L(\tau)$ i.e. every vertex has degree at most 2.

Furthermore, one has $\gamma_0 = 1$ and $\zeta \leq \gamma_k = 1 + \beta$ for $|k|_{\mathfrak{s}} = 1$. Therefore $\mathring{B}^\tau$ consist only of $\mathbf{k}$ such that $|k|_{\mathfrak{s}} = 1$ for all $k \in \mathbf{k}$. Since $D_2^2 \Upsilon^\tau = 0$, one has $\mathbf{k} \in B^\tau$ only if $|\mathbf{k}| \leq 1$ and $|\mathbf{k}|_{\mathfrak{s}} \leq 1$.

Note that $|\tau|_{\mathfrak{s}} = (L-1)(\beta+1) + \beta$. Then for $\mathbf{k} \in B^\tau$ with $|\mathbf{k}| = |\mathbf{k}|_{\mathfrak{s}} = 1$, one has

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim \vertiii{f}_\eta^L |\varphi|_{(\eta-\theta)L} \,. \tag{6.33}$$

Set $\nabla\eta(\tau, \mathbf{k}) = 0$ and $\eta(\tau, \mathbf{k}) = (\eta - \theta)L$. Then $\eta(\tau, \mathbf{k}) \geq 0$ and condition (6.12) is equivalent to

$$\Delta(\tau, \mathbf{k}) = (L-1)(\beta+1) + \beta + 1 - (\alpha\eta - \alpha - 1)L = L(\beta + 2 + \alpha - \alpha\eta) > 0 \,,$$

which holds by the assumption (6.29).

For $\mathbf{k} \in B^\tau$ with $|\mathbf{k}| \leq 1$ and $|\mathbf{k}|_{\mathfrak{s}} = 0$, one has

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})| \lesssim \vertiii{f}_\eta^L |\varphi|_{(\eta-\theta)(L-1)} (|\varphi|_{\eta-|\mathbf{k}|} + |\nabla\varphi|_1 |\varphi|_{\eta-\theta-|\mathbf{k}|}) \tag{6.34}$$

Set $\nabla\eta(\tau, \mathbf{k}) = 1$ and $\eta(\tau, \mathbf{k}) = (\eta - \theta)(L-1) + \eta - |\mathbf{k}|$. Then $\bar{\eta}(\tau, \mathbf{k}) = \eta(\tau, \mathbf{k}) - \theta \geq 0$ due to $D^{\mathbf{k}}\Upsilon^\tau \neq 0$ and the assumption that $\eta \geq \theta + 1$ if f is not a polynomial, and condition (6.12) is equivalent to

$$\Delta(\tau, \mathbf{k}) = (L-1)(\beta+1) + \beta + 2 + \alpha - \alpha|\mathbf{k}| - (\alpha\eta - \alpha - 1)(L-1) - \alpha\eta + \alpha|\mathbf{k}| = L(\beta + 2 + \alpha - \alpha\eta) > 0 \,,$$

which is the same condition as earlier and holds by (6.29). The bounds (6.33)-(6.34) imply that, with the above choices for $\eta(\tau, \mathbf{k}), \nabla\eta(\tau, \mathbf{k})$,

$$\vvvert D^{\mathbf{k}}\Upsilon^\tau \vvvert \lesssim \vertiii{f}_\eta^L \,.$$

Note also that (6.14) holds for the above choices because $|\tau|_{\mathfrak{s}} > -1$. Therefore Assumption 6.22 holds and Theorem 6.35 yields the result. ∎

Remark 6.46. Similar to Remark 6.44, under a natural generalisation of $\vertiii{f}_\eta$ in the spirit of Section 5.2, the best bound that Theorem 6.35 provides appears not to be a power of $\vertiii{f}_{\eta,q} \vvvert Z \vvvert$.

6.4 Setup for proof

In the rest of the article, we prove Theorem 6.35. As usual, we apply Corollary 3.5, so we put ourselves in the setting of Section 3. We suppose Assumptions 6.1, 6.4, 6.7, 6.16, 6.22 are in place in the rest of the paper, but remark that Assumption 6.22 is used only in Sections 6.9-6.10 and Assumption 6.4 is used only in Section 6.10.

Domains. Our basic domain is $\Sigma = \operatorname{cl}\Omega_a$ for $a \in (0,1]$ with parabolic boundary $\partial\Omega_a$, metric $(x,y) \mapsto |x-y|_{\mathfrak{s}}$, and distance to boundary $\|\!|z|\!\|_a$.

Our rescaled domains are $\Sigma_{z,\lambda} = \operatorname{cl}\Omega_b$ where $0 < b \ll 1$ will be chosen later (b and a are not related in any way). For $z \in \Omega_a$ and $0 < \lambda \le \|\!|z|\!\|_a$, the injection $T_{z,\lambda}\colon \Omega_b \to \Omega_a$ is the map from (1.4), which clearly satisfies (3.1) since $|z - T_{z,\lambda}x|_{\mathfrak{s}} \le \lambda b \le \lambda$ for all $x \in \Omega_b$.

Drivers. We let $\mathbb{M}$ denote the set of pairs (f, Z) where Z is a model on Ω_a adapted to K (Definition 6.29) and $f\colon \mathcal{J} \to L(\mathbb{R}^n, E)$ satisfies Assumption 6.22. We make the same definition for $\mathbb{M}_{z,\lambda}$ except that Z is a model on Ω_b and we require it to be adapted to the rescaled kernel

$$K^\lambda(z) = \lambda^{|\mathfrak{s}|-2} K(\lambda \cdot_{\mathfrak{s}} z) \tag{6.35}$$

(i.e. Definition 6.29 holds for $Z = (\Pi, \Gamma)$ but with K replaced by K^λ.)

Remark 6.47. In Sections 4-5, we apply Theorem 3.4 with $\mathbb{M}_{z,\lambda} = \mathbb{M}$. Here, in contrast, $\mathbb{M}_{z,\lambda} \neq \mathbb{M}$ in general. Moreover, while $\mathbb{M}_{z,\lambda}$ does not depend on z, it *does* depend on λ.

Our rescaling maps $R_{z,\lambda}\colon \mathbb{M} \to \mathbb{M}_{z,\lambda}$ are defined as follows. For $\lambda > 0$, define the map

$$Q_\lambda\colon \mathcal{J} \to \mathcal{J}\ , \qquad \nabla^k(Q_\lambda \boldsymbol{\varphi}) = \lambda^{-\alpha-|k|_{\mathfrak{s}}} \nabla^k \varphi\ , \tag{6.36}$$

which is a linear bijection. For $f\colon \mathcal{J} \to L(\mathbb{R}^n, E)$, denote

$$f_\lambda = f \circ Q_\lambda\colon \mathcal{J} \to L(\mathbb{R}^n, E)\ .$$

Recall the critical scaling $\mathfrak{c}(\tau)$ from Definition 6.10. Define the linear bijection

$$\mathfrak{C}_\lambda\colon \mathcal{T}_{\le 2} \to \mathcal{T}_{\le 2}\ , \qquad \mathfrak{C}_\lambda \tau = \lambda^{-\mathfrak{c}(\tau)} \tau$$

for all $\tau \in \mathfrak{T}_{\le 2}$. For a model $Z = (\Pi, \Gamma)$ on Ω_a, we define for $x, y \in \Omega_b$,

$$(R_{z,\lambda}\Pi)_x \tau = \big(\Pi_{T_{z,\lambda}(x)} \mathfrak{C}_\lambda \tau\big) \circ T_{z,\lambda}\ , \tag{6.37}$$

$$(R_{z,\lambda}\Gamma)_{x,y} = \mathfrak{C}_\lambda^{-1} \Gamma_{T_{z,\lambda}(x), T_{z,\lambda}(y)} \mathfrak{C}_\lambda\ , \tag{6.38}$$

and then $R_{z,\lambda}Z = (R_{z,\lambda}\Pi, R_{z,\lambda}\Gamma)$. We show in Lemma 6.48 that $R_{z,\lambda}Z$ is a model on Ω_b adapted to K^λ. We then define

$$R_{z,\lambda}\colon \mathbb{M} \to \mathbb{M}_{z,\lambda}\ , \qquad R_{z,\lambda}(f, Z) = (f_\lambda, R_{z,\lambda}Z)\ . \tag{6.39}$$

To introduce our 'norms', we define for $\lambda > 0$, $\tau \in \mathfrak{W}_{<0}$, and a model $Z = (\Pi, \Gamma)$ on Ω_b,

$$\|\Pi\|_{\tau;\lambda} = \sup_{x \in \Omega_b} \|\Pi_x \tau\|_{|\tau|_{\mathfrak{s}}; x; 1/\lambda; r}\ . \tag{6.40}$$

Note that for $\lambda \ll 1$, $\|\Pi\|_{\tau;\lambda}$ depends on Π_x evaluated on non-anticipative test functions with support of diameter order $1/\lambda \gg 1$. Then define for $z \in \Omega_a$, $0 < \lambda \le \|\!|z|\!\|$, and $(f, Z) \in \mathbb{M}_{z,\lambda}$,

$$\|(f,Z)\|_{z,\lambda} = \max_{|\tau|_{\mathfrak{s}} < 0} \max_{\mathbf{k} \in B^\tau} \Big(\|\!| D^{\mathbf{k}} \Upsilon^\tau |\!\| \|\Pi\|_{\tau;\lambda} \Big)^{\varrho(\tau,\mathbf{k})} + \max_{\mathcal{I}^k[\sigma] \in \mathfrak{U}} \Big(\|\!| \Upsilon^\sigma |\!\| \|\Gamma\|_{\mathcal{I}^k[\sigma]; \Omega_b} \Big)^{\varrho(\sigma,0)} \tag{6.41}$$

where $\varrho(\tau, \mathbf{k})$ is defined in (6.28). We define the local norms $[\![(f,Z)]\!]_{z,\lambda}$ by (6.27). Note that (3.2) clearly holds. We show that (3.3) holds in Lemma 6.51.

Solutions. For a driver $M = (f,Z) \in \mathbb{M}$, the set of solutions $\mathbb{S}_M$ is defined as $\varphi \in \mathcal{C}(\operatorname{cl}\Omega_a, E)$ that solve the remainder equation of (6.1) on Ω_a as in Definition 6.32. Likewise, for $M = (f,Z) \in \mathbb{M}_{z,\lambda}$, the set $\mathbb{S}_M$ consists of $\varphi \in \mathcal{C}(\operatorname{cl}\Omega_b, E)$ that solve the remainder equation of (6.1) on Ω_b. Closure under rescaling (3.4) is verified in Lemma 6.52.

Local coercivity. This is the hardest part of the proof and is verified in Lemma 6.70.

6.5 Scaling

Let $z \in \Omega_a$, $0 < \lambda \le |\!|\!|z|\!|\!|$, $(f,Z) \in \mathbb{M}$, and $b \in (0,1]$.

Lemma 6.48. *$R_{z,\lambda}Z$ is a model on Ω_b that is adapted to K^λ from* (6.35)*. Moreover*

$$\|R_{z,\lambda}\Gamma\|_{\mathcal{I}^k[\sigma];\Omega_b} = \lambda^{|\sigma|_{\mathfrak{s}}-\mathfrak{c}(\sigma)}\|\Gamma\|_{\mathcal{I}^k[\sigma];B_z(\lambda b)}\ , \tag{6.42}$$

$$\|R_{z,\lambda}\Pi\|_{\tau;\lambda} = \lambda^{|\tau|_{\mathfrak{s}}-\mathfrak{c}(\tau)}\|\Pi\|_{\tau;B_z(\lambda b)}\ . \tag{6.43}$$

Proof. We use the shorthands $R = R_{z,\lambda}$ and $T = T_{z,\lambda}$. We first show (6.42)-(6.43) and (a)-(c) of Definition 6.29. The algebraic identities

$$(R\Gamma)_{x,y}(R\Gamma)_{y,z} = (R\Gamma)_{x,z}\ , \qquad (R\Pi)_x(R\Gamma)_{x,y} = (R\Pi)_y$$

follow easily from the definitions, which verifies Definition 6.29(a).

For the analytic bounds of Definition 6.29(b), we have for $\mathcal{I}^k[\sigma] \in \mathfrak{U}$ and $x, y \in \Omega_b$,

$$|\langle (R\Gamma)_{x,y}\mathcal{I}^k[\sigma], \mathbf{1}\rangle| = \lambda^{-\mathfrak{c}(\mathcal{I}^k\sigma)}\langle \Gamma_{T(x),T(y)}\mathcal{I}^k[\sigma], \mathbf{1}\rangle$$

and since $|T(x) - T(y)|_{\mathfrak{s}} = \lambda|x-y|_{\mathfrak{s}}$ and $|\mathcal{I}^k\sigma|_{\mathfrak{s}} - \mathfrak{c}(\mathcal{I}^k\sigma) = |\sigma|_{\mathfrak{s}} - \mathfrak{c}(\sigma)$, we obtain (6.42). Furthermore, by (1.10),

$$\|(R\Pi)_x\tau\|_{|\tau|_{\mathfrak{s}};x;1/\lambda;r} = \lambda^{|\tau|_{\mathfrak{s}}-\mathfrak{c}(\tau)}\|\Pi_{T(x)}\tau\|_{|\tau|_{\mathfrak{s}};T(x);1;r}$$

which proves (6.43) and verifies Definition 6.29(b). To prove that RZ is a model, it remains to verify Definition 6.29(c). To this end,

$$(R\Pi)_x X^k\tau = \lambda^{-|k|_{\mathfrak{s}}}\{\Pi_{T(x)}X^k\mathfrak{C}_\lambda\tau\}\circ T = \lambda^{-|k|_{\mathfrak{s}}}(T(\cdot)-T(x))^k(R\Pi)_x\tau = (\cdot - x)^k(R\Pi)_x\tau\ , \tag{6.44}$$

which verifies the first identity in Definition 6.29(c). The second identity follows similarly.

Finally, to show adaptedness of RZ to K^λ, we need to verify (6.21). For $\ell = 0$ in (6.21),

$$(R\Pi)_x\mathcal{I}\tau = \lambda^{-\mathfrak{c}(\tau)-2}(\Pi_{T(x)}\mathcal{I}\tau)\circ T = \lambda^{-\mathfrak{c}(\tau)-2}\{K * \Pi_{T(x)}\tau - \Pi_{T(x)}\mathcal{J}(T(x))\tau\}\circ T\ , \tag{6.45}$$

where we used the definition of R in the first equality and the adaptedness of Z in the second equality. Note that, by definition of R and of K^λ in (6.35),

$$\lambda^{-\mathfrak{c}(\tau)-2}\{K * \Pi_{T(x)}\tau\}\circ T = K^\lambda * (R\Pi)_x\tau\ . \tag{6.46}$$

Moreover,

$$\begin{aligned}\lambda^{-\mathfrak{c}(\tau)-2}\{\Pi_{T(x)}\mathcal{J}(T(x))\tau\}\circ T &= \lambda^{-\mathfrak{c}(\tau)-2}\sum_{|k|_{\mathfrak{s}}<|\tau|_{\mathfrak{s}}+2}\frac{\lambda^{|k|_{\mathfrak{s}}}(\cdot - x)^k}{k!}\{D^kK * \Pi_{T(x)}\tau\}(T(x)) \\ &= \sum_{|k|_{\mathfrak{s}}<|\tau|_{\mathfrak{s}}+2}\frac{\Pi_x X^k}{k!}\{D^kK^\lambda * (R\Pi)_x\tau\}(x) \\ &= (R\Pi)_x(R\mathcal{J})(x)\tau\end{aligned} \tag{6.47}$$

where $R\mathcal{J}$ is defined as in (6.22) with Π replaced by $R\Pi$ and K replaced by K^λ, and where for the first equality we used (6.44) with $\tau = \mathbf{1}$ (so that $\Pi_x\mathbf{1} = 1$), while for the second equality we used that $\lambda^{|k|_\mathfrak{s}}\{D^k K * \Pi_{T(x)}\tau\}(T(x)) = D^k(\{K * \Pi_{T(x)}\tau\} \circ T)(x)$ followed by (6.46). Substituting (6.46)-(6.47) into (6.45) verifies (6.21) for $R\Pi$ and K^λ and $\ell = 0$. The case $\ell \neq 0$ follows from $\Pi_x\mathcal{I}^\ell\tau = D^\ell\Pi_x\mathcal{I}\tau$ and thus $(R\Pi)_x\mathcal{I}^\ell\tau = \lambda^{-\mathfrak{c}(\tau)-2+|\ell|_\mathfrak{s}}(D^\ell\Pi_{T(x)}\mathcal{I}\tau) \circ T = D^\ell(R\Pi)_x\mathcal{I}\tau$. ∎

We let Υ^τ_λ be the elementary differentials built from f_λ. Recall the map Q_λ from (6.36).

Lemma 6.49. *$\Upsilon^\tau \circ Q_\lambda = \lambda^{-2-\alpha-\mathfrak{c}(\tau)}\Upsilon^\tau_\lambda$ for all $\lambda > 0$ and conforming trees τ.*

Note that $-2-\alpha-\mathfrak{c}(\tau)$ can be both negative or positive. It is negative if f is constant because $-2-\alpha = \mathfrak{c}(\Xi_1)$ and every tree satisfies $\mathfrak{c}(\tau) \geq \mathfrak{c}(\Xi_1)$. But for non-constant f, there are conforming trees for which $\mathfrak{c}(\tau) < \mathfrak{c}(\Xi_1)$.

Proof. We proceed by induction. For the base cases, one has $\Upsilon^{\mathbf{1}}(Q_\lambda\boldsymbol{\varphi}) = P(Q_\lambda\boldsymbol{\varphi}) = \lambda^{-\alpha p}P(\boldsymbol{\varphi}) = \lambda^{-\alpha-2}\Upsilon^{\mathbf{1}}(\boldsymbol{\varphi})$ where we used $-\alpha-2 = -\alpha p$. Moreover, for $i = 1, \dots, n$, by definition $\Upsilon^{\Xi_i}(Q_\lambda\boldsymbol{\varphi}) = (f_\lambda)_i(\boldsymbol{\varphi}) = \Upsilon^{\Xi_i}_\lambda(\boldsymbol{\varphi})$, which proves the claim since $\mathfrak{c}(\Xi_i) = -\alpha-2$.

Suppose the claim is true for τ. Then, treating $i \in [d]$ canonically as an element of $\mathbb{N}^d$,

$$\begin{aligned}(\Upsilon^{X^i\tau} \circ Q_\lambda)(\boldsymbol{\varphi}) &= \sum_{\ell\in\mathbb{N}^d} \lambda^{-\alpha-|\ell|_s-|i|_\mathfrak{s}}\boldsymbol{\varphi}_{\ell+i}(D^\ell\Upsilon^\tau)(Q_\lambda\boldsymbol{\varphi}) = \sum_{\ell\in\mathbb{N}^d} \lambda^{-|i|_\mathfrak{s}}\boldsymbol{\varphi}_{\ell+i}D^\ell(\Upsilon^\tau \circ Q_\lambda)(\boldsymbol{\varphi}) \\ &= \sum_{\ell\in\mathbb{N}^d} \lambda^{-|i|_\mathfrak{s}-2-\alpha-\mathfrak{c}(\tau)}\boldsymbol{\varphi}_{\ell+i}D^\ell\Upsilon^\tau_\lambda(\boldsymbol{\varphi}) = \lambda^{-2-\alpha-\mathfrak{c}(X^i\tau)}\Upsilon^{X^i\tau}_\lambda(\boldsymbol{\varphi})\,,\end{aligned}$$

where we used the induction hypothesis in the third equality. So it suffices to consider $\tau = \Xi_j\prod_{i=1}^\ell \mathcal{I}^{k_i}[\sigma_i]$. Note that

$$\mathfrak{c}(\tau) = \mathfrak{c}(\Xi_j) + \sum_i \mathfrak{c}(\sigma_i) + 2\ell - \sum_i |k_i|_\mathfrak{s}\,. \tag{6.48}$$

Then, if $j \in \{1, \dots, n\}$,

$$\begin{aligned}\Upsilon^\tau \circ Q_\lambda(\boldsymbol{\varphi}) &= \Big[\Big(\prod_i D^{k_i}\Big)\Upsilon^{\Xi_j}\Big](Q_\lambda\boldsymbol{\varphi})(\Upsilon^{\sigma_1}(Q_\lambda\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}(Q_\lambda\boldsymbol{\varphi})) \\ &= \lambda^{-\sum_i \mathfrak{c}(\sigma_i)-2\ell-\alpha\ell}\Big[\Big(\prod_i D^{k_i}\Big)\Upsilon^{\Xi_j}\Big](Q_\lambda\boldsymbol{\varphi})(\Upsilon^{\sigma_1}_\lambda(\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}_\lambda(\boldsymbol{\varphi})) \\ &= \lambda^{-\sum_i \mathfrak{c}(\sigma_i)-2\ell-\alpha\ell+\sum_i|k_i|_\mathfrak{s}+\alpha\ell}\Big[\Big(\prod_i D^{k_i}\Big)(f_j \circ Q_\lambda)\Big](\boldsymbol{\varphi})(\Upsilon^{\sigma_1}_\lambda(\boldsymbol{\varphi}), \dots, \Upsilon^{\sigma_\ell}_\lambda(\boldsymbol{\varphi})) \\ &= \lambda^{-\mathfrak{c}(\tau)+\mathfrak{c}(\Xi_j)}\Upsilon^\tau_\lambda(\boldsymbol{\varphi}) = \lambda^{-\mathfrak{c}(\tau)-\alpha-2}\Upsilon^\tau_\lambda(\boldsymbol{\varphi})\end{aligned}$$

where we used the induction hypothesis in the second equality and (6.48) in the fourth equality. For $j = 0$, the same equalities hold except f_j is replaced by P and we use that $P \circ Q_\lambda = \lambda^{-\alpha p}P$ to conclude that the final line becomes $\lambda^{-\mathfrak{c}(\tau)-\alpha p}\Upsilon^\tau_\lambda(\boldsymbol{\varphi}) = \lambda^{-\mathfrak{c}(\tau)-\alpha-2}\Upsilon^\tau_\lambda(\boldsymbol{\varphi})$. ∎

The following lemma gives a scaling bound analogous to (5.17).

Lemma 6.50. *Under Assumption 6.22, for all $\tau \in \mathfrak{W}_{\leq 0}$ and $\mathbf{k} \in B^\tau$,*

$$\|\!| D^{\mathbf{k}}\Upsilon^\tau_\lambda |\!\| \leq \lambda^{\alpha+2+\mathfrak{c}(\tau)-|\mathbf{k}|\alpha-|\mathbf{k}|_\mathfrak{s}-\alpha\eta(\tau,\mathbf{k})}\|\!| D^{\mathbf{k}}\Upsilon^\tau |\!\|\,.$$

Proof. By Lemma 6.49, for $\lambda \in (0, 1]$,

$$D^{\mathbf{k}}\Upsilon^\tau_\lambda = \lambda^{\alpha+2+\mathfrak{c}(\tau)}D^{\mathbf{k}}(\Upsilon^\tau \circ Q_\lambda) = \lambda^{\alpha+2+\mathfrak{c}(\tau)-|\mathbf{k}|\alpha-|\mathbf{k}|_\mathfrak{s}}(D^{\mathbf{k}}\Upsilon^\tau) \circ Q_\lambda\,.$$

We conclude by (6.11) and (6.13) of Assumption 6.22 and that $|rx|_\eta \leq r^\eta|x|_\eta$ for $r \geq 1$. ∎

Lemma 6.51. $\|R_{z,\lambda}(f,Z)\|_{z,\lambda} \le \lambda^\alpha [\![(f,Z)]\!]_{z,\lambda}$.

Proof. This follows by combining Lemmas 6.48 and 6.50, definitions (6.27) and (6.41), and the fact that $b \le 1$. ∎

Lemma 6.52. *Suppose $\varphi \in \mathbb{S}_M$. Recall $R_{z,\lambda}\varphi = \lambda^\alpha \varphi \circ T_{z,\lambda}$. Then $R_{z,\lambda}\varphi \in \mathbb{S}_{R_{z,\lambda}M}$.*

Proof. Denote again $T = T_{z,\lambda}, R = R_{z,\lambda}$. We extend the bijection Q_λ from (6.36) to $\mathcal{T}_{\le 2} \to \mathcal{T}_{\le 2}$ by $Q_\lambda \tau = \lambda^{-\alpha - \mathfrak{c}(\tau)}\tau$ and correspondingly to $\mathcal{T}_{\le 2} \otimes E \to \mathcal{T}_{\le 2} \otimes E$ by scaling the left tensor.

Define $\Psi \colon \Omega_b \to \mathcal{U}_{\le 2} \otimes E$ by

$$\Psi = Q_\lambda^{-1} \Phi \circ T \,,$$

where Φ is defined as above Definition 6.32. Note that the jet $\boldsymbol{\psi} = Q_\lambda^{-1}\boldsymbol{\varphi} \circ T$ is the polynomial component of Ψ (up to degree 1) and that $\psi = \nabla^0 \psi = \lambda^\alpha \varphi \circ T = R_{z,\lambda}\varphi$. To see that Ψ is coherent in the sense of (6.25) for $\boldsymbol{\psi}$ and $R(f,Z)$, we apply Lemma 6.49 to obtain

$$\Upsilon^\tau_\lambda(\boldsymbol{\psi}) = \lambda^{2+\alpha+\mathfrak{c}(\tau)} \Upsilon^\tau(\boldsymbol{\varphi} \circ T) \,. \tag{6.49}$$

To show that Ψ is a modelled distribution, for $\sigma \in \mathfrak{T}_{\le 2}$,

$$\begin{aligned}
\langle \sigma, \Psi_x - (R\Gamma)_{xy}\Psi_y \rangle &= \Big\langle \lambda^{\alpha+\mathfrak{c}(\sigma)}\sigma, \Phi_{T(x)} - \sum_{|k|_{\mathfrak{s}} \le 2} \nabla^k \varphi_{T(y)} \Gamma_{T(x),T(y)} \frac{X^k}{k!} \\
&\qquad - \sum_{-2 < |\tau|_{\mathfrak{s}} < 0} \Upsilon^\tau(\boldsymbol{\varphi}(T(y))) \Gamma_{T(x),T(y)} \frac{\mathcal{I}[\tau]}{\tau!} \Big\rangle \\
&= \langle \lambda^{\alpha+\mathfrak{c}(\alpha)}\sigma, \Phi_{T(x)} - \Gamma_{T(x),T(y)}\Phi_{T(y)} \rangle
\end{aligned}$$

where we used in the first equality the definition of $R\Gamma$ in (6.38) and Q_λ. The final quantity is of order $\lambda^{2+\zeta-|\sigma|_{\mathfrak{s}}+\alpha+\mathfrak{c}(\alpha)}|x-y|_{\mathfrak{s}}^{2+\zeta-|\sigma|_{\mathfrak{s}}}$ since we assumed Φ is in $\mathcal{D}^{2+\zeta}(\Omega_a, E)$, therefore Ψ is in $\mathcal{D}^{2+\zeta}(\Omega_b, E)$ as desired.

Finally, we need to verify (6.26) for $\boldsymbol{\psi}$ and $R\Pi$. Observe that

$$\mathcal{L}\psi = \lambda^{\alpha+2}(\mathcal{L}\varphi) \circ T \,. \tag{6.50}$$

On the other hand, for $x \in \Omega_b$,

$$\sum_{\tau \in \mathfrak{W}_{\le 0}} \Upsilon^\tau_\lambda(\boldsymbol{\psi}_x) \frac{1}{\tau!} (R\Pi)_x \tau = \sum_{\tau \in \mathfrak{W}_{\le 0}} \lambda^{2+\alpha} \frac{1}{\tau!} \Upsilon^\tau(\boldsymbol{\varphi}_{T(x)}) \{\Pi_{T(x)}\tau\} \circ T \tag{6.51}$$

where we used the definition of $R\Pi$ in (6.37) and (6.49). It follows that

$$\begin{aligned}
&\Big\| \mathcal{L}\psi - \sum_{\tau \in \mathfrak{W}_{\le 0}} \Upsilon^\tau_\lambda(\boldsymbol{\psi}_x) \frac{1}{\tau!} (R\Pi)_x \tau \Big\|_{\zeta; x; \frac{1}{2}\|x\|_b; r} \\
&\qquad = \lambda^{\alpha+2+\zeta} \Big\| \mathcal{L}\varphi - \sum_{\tau \in \mathfrak{W}_{\le 0}} \Upsilon^\tau(\boldsymbol{\varphi}_{T(x)}) \frac{1}{\tau!} (\Pi_{T(x)}\tau) \Big\|_{\zeta; T(x); \lambda\frac{1}{2}\|x\|_b; r} < \infty \,,
\end{aligned}$$

where we used (6.50)-(6.51) and (1.10) for the equality and (6.26) and the fact that $\lambda \frac{1}{2}\|x\|_b \le \frac{1}{2}\|T(x)\|_a$ due to $\lambda \le \|z\|_a$ and $b \le 1$ for the final bound. ∎

6.6 Structure of elementary differentials

We now turn to the proof of local coercivity, which takes up the rest of the paper. The main result of this subsection is Lemma 6.53, which gives a Taylor expansion formula for elementary differentials, similar to (5.18) for rough paths.

Recall the inner product (6.5). An equivalent way to define the inner product (and thus $\tau!$) is via the recursion

$$\Big\langle X^k \Xi_j \prod_{i=1}^{\ell} \mathcal{I}^{k_i}[\tau_i], X^{\bar{k}} \Xi_{\bar{j}} \prod_{i=1}^{\ell} \mathcal{I}^{\bar{k}_i}[\bar{\tau}_i] \Big\rangle = \delta_{j,\bar{j}} \delta_{k,\bar{k}} k! \sum_{\pi \in S_\ell} \prod_{i=1}^{\ell} \delta_{k_i, \bar{k}_{\pi(i)}} \langle \tau_i, \bar{\tau}_{\pi(i)} \rangle \,, \tag{6.52}$$

where S_ℓ is the permutation group on $[\ell]$, and with $\langle \tau, \bar{\tau} \rangle = 0$ if $\tau, \bar{\tau}$ have different number of edges incident with their roots.

Lemma 6.53. *Let $\Gamma \in \mathcal{G}$, and $\tau \in \mathfrak{W}_{\le 0}$. Then*

$$\begin{aligned} &\sum_{\sigma \in \mathfrak{W}_{\le 0}} \frac{1}{\sigma!} \Upsilon^\sigma \langle \tau, \Gamma \sigma \rangle \\ &\quad = \sum_{a \in \mathbb{N}^d} \sum_{l \ge 0} \sum_{|\mathbf{k}| = l} \sum_{(\sigma_1, \dots, \sigma_l)} \frac{(\partial^a D^{\mathbf{k}} \Upsilon^\tau)}{a! \mathbf{k}!} \big(\Upsilon^{\sigma_1}, \dots, \Upsilon^{\sigma_l}\big) \Big\{ \prod_{i=1}^{l} \frac{1}{\sigma_i!} \langle \Gamma \mathcal{I}^{k_i}[\sigma_i], \mathbf{1} \rangle \Big\} \langle \Gamma X, \mathbf{1} \rangle^a \end{aligned} \tag{6.53}$$

where, on the right-hand side, the third sum is over all multi-indexes $\mathbf{k} = \{k_1, \dots, k_l\} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$ of length $l \ge 0$ and the fourth sum is over all l-tuples of trees $(\sigma_1, \dots, \sigma_l)$ in $\mathfrak{W}_{<0}$ such that $\mathcal{I}^{k_i}[\sigma_i] \in \mathfrak{U}$ and $|\tau|_{\mathfrak{s}} + \sum_{i=1}^{l} |\mathcal{I}^{k_i}[\sigma_i]|_{\mathfrak{s}} + |a|_{\mathfrak{s}} \le 0$. In particular,

$$\begin{aligned} \sum_{\sigma \in \mathfrak{W}_{\le 0}} \frac{1}{\sigma!} \Upsilon^\sigma \langle \tau, \Gamma \sigma \rangle = \sum_{l,b \ge 0} \sum_{|\mathbf{k}| = l} \sum_{|\mathbf{m}| = b} \sum_{\substack{(\sigma_1, \dots, \sigma_l) \\ (a_1, \dots, a_b)}} \frac{D^{\mathbf{k}} D^{\mathbf{m}} \Upsilon^\tau}{\mathbf{k}! \mathbf{m}!} \big(\Upsilon^{\sigma_1}, \dots, \Upsilon^{\sigma_l}, \nabla^{m_1 + a_1}, \dots, \nabla^{m_b + a_b}\big) \\ \times \Big\{ \prod_{i=1}^{l} \frac{1}{\sigma_i!} \langle \Gamma \mathcal{I}^{k_i} \sigma_i, \mathbf{1} \rangle \Big\} \Big\{ \prod_{i=1}^{b} \frac{1}{a_i!} \langle \Gamma X, \mathbf{1} \rangle^{a_i} \Big\} \end{aligned} \tag{6.54}$$

where $\mathbf{m}$ runs over all multi-indexes $\mathbf{m} = \{m_1, \dots, m_b\} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$ of length b, $(a_1, \dots, a_b)$ runs over all b-tuples with $a_i \in \mathbb{N}^d$ and $a_i \neq 0$, and where we again restrict the sum to $|\tau|_{\mathfrak{s}} + \sum_{i=1}^{l} |\mathcal{I}^{k_i}[\sigma_i]|_{\mathfrak{s}} + \sum_{i=1}^{b} |a_i|_{\mathfrak{s}} \le 0$. Above, we write ∇^k for the projection map $\nabla^k \colon \boldsymbol{\varphi} \mapsto \nabla^k \varphi$.

Remark 6.54. Recall the notation from Definition 6.20. For every $\mathbf{k}, \mathbf{m}$ in (6.54), $\gamma_{\mathbf{m}} \le |\mathbf{m}| = b \le \sum_{i=1}^{b} |a_i|_{\mathfrak{s}}$, and, by definition of γ_k in (6.9), $\gamma_{\mathbf{k}} \le \sum_{i=1}^{l} |\mathcal{I}^{k_i}[\sigma_i]|_{\mathfrak{s}}$. Therefore $\gamma_{\mathbf{k}+\mathbf{m}} \le -|\tau|_{\mathfrak{s}}$ and thus either $\mathbf{k} + \mathbf{m} \in \mathring{B}^\tau$ or $D^{\mathbf{k}} D^{\mathbf{m}} \Upsilon^\tau = 0$. In particular, the right-hand side of (6.54) is well-defined for f satisfying the regularity condition in Assumption 6.22 (similar considerations apply to the right-hand side of (6.53)).

To simplify the proof below, we assume f is smooth so that Υ^τ is well-defined for any tree τ, but by continuity, the final equalities (6.53)-(6.54) hold for any f satisfying Assumption 6.22.

Proof of (6.53) $\Rightarrow$ (6.54)**.** Observe that, for any $\mathbf{t} \in \mathbb{R}^d$ and smooth function F of $\boldsymbol{\varphi}$,

$$\sum_{a \in \mathbb{N}^d} \frac{\mathbf{t}^a}{a!} \partial^a F(\boldsymbol{\varphi}) = \sum_{\mathbf{m} \in \mathbb{N}^{\mathbb{N}^d}} \frac{D^{\mathbf{m}} F}{\mathbf{m}!}(\boldsymbol{\varphi}) \Big(\sum_{a \neq 0} \frac{\mathbf{t}^a}{a!} \partial^a \boldsymbol{\varphi} \Big)^{\mathbf{m}} \tag{6.55}$$

as an equality between formal power series, where $\partial^a \boldsymbol{\varphi}$ is the jet defined by $[\partial^a(\boldsymbol{\varphi})]_j = \nabla^{a+j} \varphi$. For $m \in \mathbb{N}^d$, note that $(\sum_{a \neq 0} \frac{\mathbf{t}^a}{a!} \partial^a \boldsymbol{\varphi})^m = \sum_{a \neq 0} \frac{\mathbf{t}^a}{a!} \nabla^{m+a} \varphi$. We then obtain (6.54) by applying (6.55) to a fixed $l \ge 0, |\mathbf{k}| = l, \sigma_1, \dots, \sigma_l$ in (6.53) and taking $\mathbf{t} = \langle \Gamma X, \mathbf{1} \rangle$ and $F(\boldsymbol{\varphi}) =$

$D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})(\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_l})$ (where we treat $\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_l}$ as fixed vectors not depending on $\boldsymbol{\varphi}$). The fact that only multi-indexes $\mathbf{m}\in\mathbb{N}^{\mathbb{N}^d_{<2}}$ remain in the sum is due to the fact that $D^m\Upsilon^\tau=0$ for all $|m|_{\mathfrak{s}}\geq 2$ and $\tau\in\mathfrak{W}_{\leq 0}$ by Lemma 6.26. ∎

To conclude the proof of Lemma 6.53, it remains to prove (6.53). To this end, it is convenient to introduce several algebraic structures stemming from pre-Lie algebras. The starting point is the observation from [BCCH21] that $\tau\mapsto\Upsilon^\tau$ is a morphism for (multi-)pre-Lie algebra structures on trees and non-linearities (see also [BCFP19] for the case of rough paths). The key point is to introduce maps M and M^+ below that generalise the pre-Lie products from [BCCH21] and interact well with the map Υ.

Definition 6.55. Let $\mathfrak{V}$ denote the set of trees (in the sense of Definition 6.8) and $\mathcal{V}$ denote the vector space spanned by $\mathfrak{V}$. Let $\mathcal{H}$ denote the free commutative algebra generated by the trees X^i with $i\in[d]$ and by trees of the form $\mathcal{I}^k\sigma$ for $\sigma\in\mathfrak{V}$ and $k\in\mathbb{N}^d$. We denote the 'empty' product by $\mathbf{1}^+$ which is the unit under multiplication. A basis for $\mathcal{H}$ is given by the set of trees of the form $X^a\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$. We also let $\mathcal{H}^*$ denote the space of series of trees in $\mathcal{H}$. We identify $\mathcal{H}^*$ with the dual space of $\mathcal{H}$ via the inner product $\langle\cdot,\cdot\rangle$.

For trees σ,τ, define $\sigma\curvearrowright_k\tau$ as the sum of trees obtained by grafting σ onto the vertices of τ with a new edge such that, if the grafting happens at vertex v, then we sum over the trees obtained by reducing the polynomial label at v to $\mathfrak{n}(v)-q$ and setting the derivative label of the new grafted edge to $k-q$ for all $q\leq\mathfrak{n}(v)$ and then multiplying the resulting tree by $\binom{k}{q}$ – see [BCCH21, Remark 4.12] or [Che22, Eq. (4.26)] for a precise formula. In addition, for $i\in[d]$, define the raising operator $\uparrow_i\tau$ as the sum of trees obtained by increasing the degree of every polynomial label by i.

We now define several maps that are related to the dual of the map Δ from Section 6.1.5. These maps are closely related to those in [BM23]; see also [BK23] for a post-Lie algebra perspective. For completeness, we give all necessary definitions.

We define a bilinear map $M^+\colon\mathcal{H}^*\times\mathcal{H}^*\to\mathcal{H}^*$ as follows. First, we set

$$M^+(\tau,\mathbf{1}^+)=0\,,\qquad M^+(\mathbf{1}^+,\tau)=\tau\,,\qquad M^+(X^j,X^i)=0\,,\qquad M^+(\mathcal{I}^k[\sigma],X^i)=0\,,$$
$$M^+(\mathcal{I}^k[\sigma],\mathcal{I}^\ell[\tau])=\mathcal{I}^\ell[\sigma\curvearrowright_k\tau]\,,\qquad M^+(X^i,\mathcal{I}^\ell[\tau])=\mathcal{I}^\ell[\uparrow_i\tau]\,.$$

We then extend $M^+(X^i,\cdot)$ and $M^+(\mathcal{I}^k[\sigma],\cdot)$ as derivations on $\mathcal{H}$ for the commutative product in $X^a\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$. We then define inductively, for $\sigma=\mathcal{I}^{k_2}[\sigma_2]\cdots\mathcal{I}^{k_r}[\sigma_r]$,

$$M^+(\mathcal{I}^{k_1}[\sigma_1]\sigma,\tau)=M^+(\mathcal{I}^{k_1}[\sigma_1],M^+(\sigma,\tau))-M^+(M^+(\mathcal{I}^{k_1}[\sigma_1],\sigma),\tau)\,,$$

which one can interpret as a sum of all possible graftings of the trees $\sigma_1,\dots,\sigma_k$ using the edges $\mathcal{I}^{k_1},\dots,\mathcal{I}^{k_r}$ onto τ (changing polynomial and derivative labels as earlier). It is easy to verify that $M^+(\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r],\tau)$ is independent of the order of terms in $\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$.

Finally, for $\sigma=\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$ and $a\neq 0$, we take any $i\in a$ and define inductively

$$M^+(X^a\sigma,\tau)=M^+(X^i,M^+(X^{a-i}\sigma,\tau))-M^+(M^+(X^i,X^{a-i}\sigma),\tau)\,,$$

which one can interpret as raising the polynomial labels in τ by a *after* grafting $\sigma_1,\dots,\sigma_k$ with the corresponding edges. Again, it is easy to verify that $M^+(X^a\sigma,\tau)$ is independent of the choice of $i\in a$ used in the inductive definition. These definitions extend canonically to the space of series $M\colon\mathcal{H}^*\times\mathcal{H}^*\to\mathcal{H}^*$ since the operations are locally finite.

We further define a bilinear map $M\colon\mathcal{H}^*\times\mathcal{V}\to\mathcal{V}$ as follows. First, we set

$$M(\mathbf{1}^+,\tau)=\tau\,,\qquad M(\mathcal{I}^k[\sigma],\tau)=\sigma\curvearrowright_k\tau\,,\qquad M(X^i,\tau)=\uparrow_i\tau\,.\tag{6.56}$$

Then we define inductively for $\sigma = \mathcal{I}^{k_2}[\sigma_2]\cdots\mathcal{I}^{k_r}[\sigma_r]$

$$M(\mathcal{I}^{k_1}[\sigma_1]\sigma,\tau) = M(\mathcal{I}^{k_1}[\sigma_1], M(\sigma,\tau)) - M(M^+(\mathcal{I}^{k_1}[\sigma_1],\sigma),\tau) \tag{6.57}$$

and, for $\sigma = \mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$ and $a \neq 0$ and any $i \in a$,

$$M(X^a\sigma,\tau) = M(X^i, M(X^{a-i}\sigma,\tau)) - M(M^+(X^i, X^{a-i}\sigma),\tau)\,. \tag{6.58}$$

As earlier, $M(X^a\mathcal{I}_1^k[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r],\tau)$ can be written as a weighted sum of trees formed by grafting $\sigma_1,\dots,\sigma_r$ onto τ using edges $\mathcal{I}^{k_1},\dots\mathcal{I}^{k_r}$ and then raising the polynomial labels of the vertices of τ by a. These operations are locally finite so extend to the space of series as before.

Lemma 6.56. *For $\tau \in \mathfrak{V}$ and $X^a\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r \in \mathcal{H}$,*

$$\Upsilon^{M(X^a\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r,\tau)} = (\partial^a D^{k_1}\cdots D^{k_r}\Upsilon^\tau)(\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_r})\,. \tag{6.59}$$

Proof. Recall the pre-Lie morphism property of Υ, see, e.g. [BCCH21, Cor. 4.15] or [Che22, Eq. (4.29)],

$$\Upsilon^{\sigma\curvearrowright_k\tau} = (D^k\Upsilon^\tau)(\Upsilon^\sigma)\,. \tag{6.60}$$

Moreover, by [BCCH21, Cor. 4.15] or [Che22, Eq. (4.30)],

$$\Upsilon^{\uparrow_i\tau} = \partial^i\Upsilon^\tau\,. \tag{6.61}$$

Then (6.59) for $r = 0$, and for $r = 1$ and $a = 0$, follows from these identities.

Proceeding by induction on r first, it follows from the definition of M that

$$\Upsilon^{M(\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r,\tau)} = \Upsilon^{M(\mathcal{I}^{k_1}\sigma_r, M(\mathcal{I}^{k_2}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_{r-1},\tau))} - \Upsilon^{M(M^+(\mathcal{I}^{k_1}\sigma_1,\mathcal{I}^{k_2}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r),\tau)}$$

The first term on the right-hand side is, by the induction hypothesis and (6.60), equal to

$$D^{k_1}\{(D^{k_2}\cdots D^{k_r}\Upsilon^\tau)(\Upsilon^{\sigma_2},\cdots,\Upsilon^{\sigma_r})\}(\Upsilon^{\sigma_1})\,,$$

while the second term, by the derivation property of $M^+(\mathcal{I}^{k_1}\sigma_1,\cdot)$ and the induction hypothesis, is equal to

$$-\sum_{i=2}^r D^{k_2}\cdots D^{k_r}\Upsilon^\tau(\Upsilon^{\sigma_2},\dots,D^{k_1}\Upsilon^{\sigma_i}(\Upsilon^{\sigma_1}),\dots,\Upsilon^{\sigma_r})\,.$$

Combining the two terms and applying the chain rule, it follows that

$$\Upsilon^{M(\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r,\tau)} = (D^{k_1}\cdots D^{k_r}\Upsilon^\tau)(\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_r})$$

as desired. Now we proceed by induction on a and write, for some $i \in a$,

$$\Upsilon^{M(X^a\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r,\tau)} = \Upsilon^{M(X^i, M(X^{a-i}\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r,\tau))} - \Upsilon^{M(M^+(X^i,X^{a-i}\mathcal{I}^{k_1}\sigma_1\cdots\mathcal{I}^{k_r}\sigma_r),\tau)}\,.$$

The first term, by the induction hypothesis and (6.61), is equal to

$$\partial^i\{(\partial^{a-i}D^{k_1}\cdots D^{k_r}\Upsilon^\tau)(\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_k})\}\,,$$

while the second term, by the induction hypothesis and the derivation property of $M^+(X^i,\cdot)$, is equal to

$$-\sum_{j=1}^r(\partial^{a-i}D^{k_1}\cdots D^{k_r}\Upsilon^\tau)(\Upsilon^{\sigma_1},\dots,\partial^i\Upsilon^{\sigma_j},\dots,\Upsilon^{\sigma_r})\,.$$

Again by the chain rule (which implies ∂^i acts as a derivation), we obtain (6.59). ∎

Recall the subalgebra $\mathcal{H}^+ \subset \mathcal{H}$ and the linear map $\Delta\colon \mathcal{T}_{\leq 2} \to \mathcal{H}^+ \otimes \mathcal{T}_{\leq 2}$ from Section 6.1.5. Let $\mathcal{T}^*_{\leq 2}$ be a copy of $\mathcal{T}_{\leq 2}$, which we identify with the dual of $\mathcal{T}_{\leq 2}$ via the inner product $\langle \cdot, \cdot \rangle$. The following result is similar to [BM23, Prop. 3.17]. For completeness, we give a direct proof.

Lemma 6.57. *For trees $\sigma \in \mathcal{H}^+$, $\tau \in \mathcal{T}^*_{\leq 2}$, and $\eta \in \mathcal{T}_{\leq 2}$,*

$$\langle M(\sigma, \tau), \eta \rangle = \langle \sigma \otimes \tau, \Delta\eta \rangle \; . \tag{6.62}$$

Proof. It is not difficult to verify (see [BCCH21, p. 939] or [Che22, Eq. (4.37)-(4.38)]) that

$$\langle \mathcal{I}^k[\sigma] \otimes \tau, \Delta\eta \rangle = \langle \sigma \curvearrowright_k \tau, \eta \rangle \; , \qquad \langle X^i \otimes \tau, \Delta\eta \rangle = \langle \uparrow_i \tau, \eta \rangle \tag{6.63}$$

for $i \in [d]$. By the base case definition of M (6.56), this proves (6.62) for $\sigma = \mathcal{I}^k[\sigma], X^i \in \mathcal{H}^+$.

Define now a coproduct $\Delta^+\colon \mathcal{H}^+ \to \mathcal{H}^+ \otimes \mathcal{H}^+$ by

$$\Delta^+ \mathbf{1}^+ = \mathbf{1}^+ \otimes \mathbf{1}^+ \; , \qquad \Delta^+ X^i = X^i \otimes \mathbf{1}^+ + \mathbf{1}^+ \otimes X^i$$

for $i \in [d]$ and

$$\Delta^+ \mathcal{I}^k[\sigma] = (\mathrm{Id} \otimes \mathcal{I}^k)\Delta\sigma + \sum_{|\ell|_{\mathfrak{s}} < |\mathcal{I}^k[\sigma]|_{\mathfrak{s}}} \mathcal{I}^{k+\ell}[\sigma] \otimes \frac{X^\ell}{\ell!} \; ,$$

and extended multiplicatively to $\mathcal{H}^+$. Using the shorthand $(\sigma \otimes \bar{\sigma})\Delta^+$ for the linear functional $\langle \sigma \otimes \bar{\sigma}, \Delta^+ \cdot \rangle\colon \mathcal{H}^+ \to \mathbb{R}$, and denoting $\tilde{\sigma} = \mathcal{I}^{k_2}[\sigma_2] \cdots \mathcal{I}^{k_r}[\sigma_r]$, we claim that

$$(\mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma})\Delta^+ = M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma}) + \mathcal{I}^{k_1}[\sigma_1]\tilde{\sigma} \; . \tag{6.64}$$

Indeed, for $\omega = \mathcal{I}^{k_1}[\sigma_1]\tilde{\sigma}$,

$$\langle \mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma}, \Delta^+ \omega \rangle = \sigma_1! \beta_1 \langle \tilde{\sigma}, \tilde{\sigma} \rangle = \langle \mathcal{I}^{k_1}[\sigma_1]\tilde{\sigma}, \omega \rangle \; ,$$

where β_1 is the number of times $\mathcal{I}^{k_1}[\sigma_1]$ appears in $\mathcal{I}^{k_1}[\sigma_1], \ldots, \mathcal{I}^{k_r}[\sigma_r]$ and where in the final equality we used the definition of the tree factorial and inner product in (6.4)-(6.5).

On the other hand, for ω of the form $\omega = \mathcal{I}^{k_2}[\omega_2] \cdots \mathcal{I}^{k_r}[\omega_r]$,

$$\begin{aligned}
\langle \mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma}, \Delta^+ \omega \rangle &= \langle \mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma}, (\mathrm{Id} \otimes \mathcal{I}^{k_2})[\Delta\omega_2] \cdots (\mathrm{Id} \otimes \mathcal{I}^{k_r})[\Delta\omega_r] \rangle \\
&= \sum_{i=2}^r \langle \mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma}, (\mathrm{Id} \otimes \mathcal{I}^{k_i})[\Delta\omega_i] \prod_{j \neq i} (\mathbf{1}^+ \otimes \mathcal{I}^{k_j}[\omega_j]) \rangle \\
&= \sum_{i=2}^r \sum_{b=2}^r \langle \mathcal{I}^{k_b}[\sigma_1 \curvearrowright_{k_1} \sigma_b], \mathcal{I}^{k_i}[\omega_i] \rangle \Big\langle \prod_{\ell \neq b} \mathcal{I}^{k_\ell}[\sigma_\ell], \prod_{j \neq i} \mathcal{I}^{k_j}[\omega_j] \Big\rangle \\
&= \sum_{b=2}^r \langle \mathcal{I}^{k_b}[\sigma_1 \curvearrowright_{k_1} \sigma_b] \prod_{\ell \neq b} \mathcal{I}^{k_\ell}[\sigma_\ell], \omega \rangle = \langle M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma}), \omega \rangle
\end{aligned}$$

where in the third equality we used the inductive definition of the inner product (6.52) and $\langle \mathcal{I}^{k_1}[\sigma_1] \otimes \mathcal{I}^{k_j}[\sigma_j], (\mathrm{Id} \otimes \mathcal{I}^{k_i})[\Delta\omega_i] \rangle = \langle \mathcal{I}^{k_j}[\sigma_1 \curvearrowright_{k_1} \sigma_j], \mathcal{I}^{k_i}[\omega_i] \rangle$, and in the fourth equality we used again (6.52). For ω not of the form above, $\langle \mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma}, \Delta^+ \omega \rangle = 0$. This proves (6.64).

Writing $\tilde{\sigma} = \mathcal{I}^{k_2}[\sigma_2] \cdots \mathcal{I}^{k_r}[\sigma_r]$, we now prove (6.62) for $\sigma = \mathcal{I}^{k_1}\tilde{\sigma}$ by induction on r. As elements of the dual of $\mathcal{T}_{\leq 2}$,

$$\begin{aligned}
(\mathcal{I}^{k_1}[\sigma_1]\tilde{\sigma} \otimes \tau)\Delta &= (\{(\mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma})\Delta^+ - M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma})\} \otimes \tau)\Delta \\
&= (\mathcal{I}^{k_1}[\sigma_1] \otimes \tilde{\sigma} \otimes \tau)(\mathrm{Id} \otimes \Delta)\Delta - (M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma}) \otimes \tau)\Delta \\
&= \{\mathcal{I}^{k_1}[\sigma_1] \otimes M(\tilde{\sigma}, \tau)\}\Delta - M(M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma}), \tau) \\
&= M(\mathcal{I}^{k_1}[\sigma_1], M(\tilde{\sigma}, \tau)) - M(M^+(\mathcal{I}^{k_1}[\sigma_1], \tilde{\sigma}), \tau) = M(\mathcal{I}^{k_1}[\sigma_1]\tilde{\sigma}, \tau)
\end{aligned} \tag{6.65}$$

where we used (6.64) in the first equality, the coassociativity property $(\mathrm{Id}\otimes\Delta)\Delta = (\Delta^+\otimes\mathrm{Id})\Delta$ (see [Hai14, Thm. 8.16]) in the second equality, induction in r in the third equality (note for the second term that $M^+(\mathcal{I}^{k_1}[\sigma_1],\tilde\sigma)$ is a linear combination of trees of the form $\mathcal{I}^{k_2}[\eta_2]\cdots\mathcal{I}^{k_r}[\eta_r]$ for which we can apply the induction hypothesis), (6.63) and (6.56) in the fourth equality, and (6.57) in the fifth equality. This concludes the proof of (6.62) for $\sigma = \mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$.

The proof for $\sigma = X^a\tilde\sigma$ where $\tilde\sigma = \mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$ and $a\neq 0$ is similar. Indeed, we have

$$(X^i\otimes X^{a-i}\tilde\sigma)\Delta^+ = M^+(X^i, X^{a-i}\tilde\sigma) + X^a\tilde\sigma\;,$$

the proof of which is similar to that of (6.64) except that we consider the two cases $\omega = X^a\tilde\sigma$ and $\omega = X^{a-i}\mathcal{I}^{k_1}[\omega_1]\cdots\mathcal{I}^{k_r}[\omega_r]$. The proof of (6.62) for $\sigma = X^a\mathcal{I}^{k_1}[\sigma_1]\cdots\mathcal{I}^{k_r}[\sigma_r]$ is then identical to (6.65) except we use induction in a instead of r and (6.58) instead of (6.57). ∎

We are finally ready for the

Proof of (6.53)**.** Define the character $g\in\mathcal{H}^+\to\mathbb{R}$ by (6.19), so that $\Gamma = (g\otimes\mathrm{Id})\Delta$. We treat $g\in\mathcal{H}^*$ by $\langle g,\sigma\rangle = g(\sigma)$ if $\sigma\in\mathcal{H}^+$ and $\langle g,\sigma\rangle = 0$ if σ is a tree that is not in $\mathcal{H}^+$. Let $\Gamma^*\colon\mathcal{T}^*_{\le 2}\to\mathcal{T}^*_{\le 2}$ be the dual map defined by the identity, for all $\sigma\in\mathcal{T}_{\le 2}$,

$$\langle\Gamma^*\tau,\sigma\rangle = \langle\tau,\Gamma\sigma\rangle = \langle g\otimes\tau,\Delta\sigma\rangle\;.$$

Then, by Lemma 6.57, for $\sigma\in\mathcal{T}_{\le 2}$,

$$\langle\Gamma^*\tau,\sigma\rangle = \langle M(g,\tau),\sigma\rangle\;.$$

Let $\pi_{\le 0}\colon\mathcal{V}\to\mathcal{V}$ denote the linear map such that, for $\sigma\in\mathfrak{V}$, $\pi_{\le 0}\sigma = \sigma$ if $\sigma\in\mathfrak{W}_{\le 0}$ and $\pi_{\le 0}\sigma = 0$ otherwise. Recall from Remark 6.15 that $\Upsilon^\sigma = 0$ if σ is not conforming. Moreover, if σ is conforming and has a polynomial leaf, then $|\sigma|_{\mathfrak{s}} > 0$ by Lemma 6.37. Recalling the definition of $\mathcal{T}_{\le 2}$ from Definition 6.11, it follows that $\Upsilon^{\pi_{\le 0}\Gamma^*\tau} = \Upsilon^{\pi_{\le 0}M(g,\tau)}$ and thus

$$\Upsilon^{\pi_{\le 0}\Gamma^*\tau} = \sum_\sigma \Upsilon^{M(\sigma,\tau)}\frac{1}{\sigma!}\langle g,X\rangle^a\prod_{i=1}^l\langle g,\mathcal{I}^{k_i}\sigma_i\rangle$$

where the sum is over all trees $\sigma\in\mathcal{H}^+$ of the form $\sigma = X^a\prod_{i=1}^l\mathcal{I}^{k_i}\sigma_i$ such that $|\sigma|_{\mathfrak{s}} + |\tau|_{\mathfrak{s}}\le 0$. It follows from Lemma 6.56 that

$$\Upsilon^{\pi_{\le 0}\Gamma^*\tau} = \sum_{a\in\mathbb{N}^d}\sum_{l\ge 0}\frac{1}{a!l!}\sum_{(\sigma_1,\dots,\sigma_l)}\sum_{(k_1,\dots,k_l)}(\partial^a D^{k_1}\cdots D^{k_l}\Upsilon^\tau)(\Upsilon^{\sigma_1},\dots,\Upsilon^{\sigma_l})\langle g,X\rangle^a\prod_{i=1}^l\frac{1}{\sigma_i!}\langle g,\mathcal{I}^{k_i}\sigma_i\rangle$$

where the final two sums are over all tuples $(\sigma_1,\dots,\sigma_l)$ and $(k_1,\dots,k_l)$ such that $\mathcal{I}^{k_i}\sigma_i\in\mathfrak{U}$ and $|\tau|_{\mathfrak{s}} + \sum_{i=1}^l|\mathcal{I}^{k_i}\sigma_i|_{\mathfrak{s}} + |a|_{\mathfrak{s}}\le 0$. Above, we used (6.59), the definition of the tree factorial (6.4), and the fact that every tree $\eta = X^a\prod_{i=1}^r(\mathcal{I}^{q_i}[\eta_i])^{m_i}$ with distinct (q_i,η_i) appears $\binom{l}{m_1,\dots,m_r}$ times in the sum over the tuples σ_j,k_j.

We now obtain (6.53) by the relation (6.19) and remarking that the left-hand side of (6.53) is precisely $\Upsilon^{\pi_{\le 0}\Gamma^*\tau}$ due to the definition of the inner product, and by using the fact that every tuple $(k_1,\dots,k_l)$ appears $\frac{l!}{\mathbf{k}!}$ times in the sum where $\mathbf{k} = \{k_1,\dots,k_l\}\in\mathbb{N}^{\mathbb{N}^d}$. ∎

The following lemma is a useful consequence of (6.59).

Lemma 6.58. *Consider $\mathbf{k}\in\mathbb{N}^{\mathbb{N}^d}$ and a conforming tree τ such that $D^{\mathbf{k}}\Upsilon^\tau\neq 0$. Then*

$$|\tau|_{\mathfrak{s}} + |\mathbf{k}|(\beta+2) - |\mathbf{k}|_{\mathfrak{s}}\ge\beta\;, \tag{6.66}$$

with equality if and only if $\mathbf{k} = 0$ and $\tau = \Xi_j$ for some $j>0$, and where we recall $\beta = |\Xi_1|_{\mathfrak{s}}$.

Proof. Since $D^{\mathbf{k}}\Upsilon^\tau \neq 0$, one has $D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi})(v,\dots,v) \neq 0$ for some vector $v \in E$ and $\boldsymbol{\varphi} \in \widehat{\mathcal{J}}$ (recall (6.6)). We include another noise label Ξ_{n+1} with corresponding function $f_{n+1} \equiv v$. Then $M(\prod_{k\in\mathbf{k}} \mathcal{I}^k[\Xi_{n+1}], \tau)$ is a sum of trees σ with $|\sigma|_{\mathfrak{s}}$ equal to the left-hand side of (6.66). Moreover, by (6.59), we have $\Upsilon^\sigma \neq 0$ for one of these trees σ, and therefore σ is conforming. The bound (6.66) therefore follows from Lemma 6.18. Moreover, Lemma 6.18 implies that $|\sigma|_{\mathfrak{s}} = \beta$ if and only if $\sigma = \Xi_j$ for some $j > 0$, but this happens if and only if $\tau = \Xi_j$ and $\mathbf{k} = 0$. ∎

6.7 Local reconstruction bounds

In this subsection, we introduce singular modelled distribution spaces and reconstruction bounds that we use later. Recall the regularity structure $(\mathcal{T}_{\leq 2}, \mathcal{G}, \mathcal{A})$ from Definition 6.29. For $\mathfrak{r} > -2$, a *sector of regularity* $\mathfrak{r}$ is a graded subspace $\mathcal{V} = \bigoplus_{a\in\mathcal{A}} \mathcal{V}_a$ such that $\mathcal{V}_a \subset \operatorname{span}\{\tau \in \mathfrak{T}_{\leq 2} : |\tau|_{\mathfrak{s}} = a\}$, $\mathcal{V}_a = \{0\}$ for $a < \mathfrak{r}$, and $\Gamma\mathcal{V} \subset \mathcal{V}$ for all $\Gamma \in \mathcal{G}$. We write $\mathcal{V}_{<\omega} = \bigoplus_{a<\omega} \mathcal{V}_a$.

Notation 6.59. In the rest of the section, we fix $0 < b \leq 1$ and write $\|\!|x|\!\| = \|\!|x|\!\|_b$ for $x \in \Omega_b$. We let τ be a tree in $\mathfrak{T}_{\leq 2}$.

Consider $\omega > 0$,[3] $\eta \in \mathbb{R}$, and a sector $\mathcal{V}$ of regularity $\mathfrak{r} \leq 0$. Fix a model $Z = (\Pi, \Gamma)$ on the open parabolic ball Ω_b. For a function $F\colon \Omega_b \to \mathcal{V}_{<\omega}$ we write

$$F_x = \sum_{|\tau|_{\mathfrak{s}}<\omega} F_x^\tau \tau \, .$$

Define the (extended) norms[4]

$$\|F\|_{\omega,\eta} = \sup_{|\tau|_{\mathfrak{s}}<\omega} \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}} \, , \quad \text{where } \|F^\tau\|_{\infty;\theta} = \sup_{x\in\Omega_b} (1 \wedge |\!|\!|x|\!|\!|^{-\theta})|F_x^\tau| \, , \tag{6.67}$$

and semi-norm

$$|F|_{\omega,\eta} = \sup_{x\in\Omega_b} \sup_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\tau,\omega,\eta} \, ,$$

where we define the two parameter function $R^{\tau,F}\colon \Omega_b^2 \to \mathbb{R}$ by

$$\sum_{|\tau|_{\mathfrak{s}}<\omega} R_{x,y}^{\tau,F}\tau = F_y - \Gamma_{yx}F_x$$

i.e. $R_{x,y}^{\tau,F} = \frac{1}{\tau!}\langle F_y - \Gamma_{yx}F_x, \tau\rangle$, and the norm

$$|R|_{\tau,\omega,\eta} = \sup_{x\in\Omega_b} |\!|\!|x|\!|\!|^{\omega-\eta} |R|_{\omega-|\tau|_{\mathfrak{s}};B_x(\frac{1}{2}|\!|\!|x|\!|\!|)} \, , \quad \text{where } |R|_{\theta;\mathfrak{K}} = \sup_{y,z\in\mathfrak{K}} \frac{|R_{y,z}|}{|y-z|_{\mathfrak{s}}^\theta} \, . \tag{6.68}$$

Let $\mathcal{D}^{\omega,\eta}(\Omega_b, \mathcal{V})$ denote the space of $F\colon \Omega_b \to \mathcal{V}_{<\omega}$ for which $\|F\|_{\omega,\eta} + |F|_{\omega,\eta} < \infty$. (This is essentially the space $\mathcal{D}^{\omega,w}$ from [GH19] with $w = (\eta,\eta,\eta)$.) We write $\mathcal{D}_{\mathfrak{r}}^{\omega,\eta}(\Omega_b)$ for $\mathcal{D}^{\omega,\eta}(\Omega_b, \mathcal{V})$ whenever $\mathcal{V}$ is clear from the context.

It follows from the local reconstruction theorem [ZK23, Thm. 1.1] (see also [Hai14, Lem. 6.7] for a less precise but simpler statement) that there is a unique distribution $\tilde{\mathcal{R}}F \in \mathcal{D}'(\Omega_b)$ such that, uniformly in $x \in \Omega_b$,

$$\|\tilde{\mathcal{R}}F - \Pi_x F_x\|_{\omega;x;\leq\frac{1}{8}|\!|\!|x|\!|\!|;r} \lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\omega-|\tau|_{\mathfrak{s}};B_x(\frac{1}{2}|\!|\!|x|\!|\!|)} \|\Pi\|_{\tau;\Omega_b} \tag{6.69}$$

where r is from Definition 6.29 and we recall the semi-norms (1.8) and (6.20). The bound (6.69) is the analogue of (5.21). The fact that $\|\Pi\|_{\tau;\Omega_b}$ appears on the right-hand side, and not $\|\Pi\|_{\tau;\mathfrak{K}}$ for some enlargement $\mathfrak{K}$ of Ω_b follows from the proof of [CCHS22, Lem. A.2].

[3] we take $\omega = \zeta$ or $\omega = 2 + \zeta$ in later subsections

[4] One should not confuse the norm $\|F^\tau\|_{\infty;\theta}$ in this section with $\|f\|_{\infty;\eta}$ from (4.15). The latter is used in Sections 4-5 and does not appear in this section.

Remark 6.60. We can localise Π around x as well, but this would not improve the final result.

Lemma 6.61. *Let $\eta \le \mathfrak{r} \le 0 < \omega$ with $\eta \notin \mathbb{Z}$ and $\lfloor 1-\eta \rfloor \ge r$ and $F \in \mathcal{D}^{\omega,\eta}_{\mathfrak{r}}(\Omega_b)$. Then there exists $\mathcal{R}F \in \mathcal{D}'(\mathbb{R}^d)$ such that $\mathrm{supp}(\mathcal{R}F) \subset \mathrm{cl}\,\Omega_b$, $\langle \mathcal{R}F, \psi\rangle = \langle \tilde{\mathcal{R}}F, \psi\rangle$ for all $\psi \in \mathcal{C}^\infty_c(\Omega_b)$, and*

$$\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)} \lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} \|\Pi\|_{\tau;\Omega_b}\{|R^{\tau,F}|_{\tau,\omega,\eta} + \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\}\ ,$$

where the proportionality constant is uniform in F, Z, b.

Proof. It follows from (6.69) that

$$\sup_{x\in\Omega_b} \|\!|x|\!\|^{\omega-\eta} \|\tilde{\mathcal{R}}F - \Pi_x F_x\|_{\omega;x;\le\frac{1}{2}\|\!|x|\!\|;r} \lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\tau,\omega,\eta}\|\Pi\|_{\tau;\Omega_b}\ . \tag{6.70}$$

(Note $\le \frac{1}{2}\|\!|x|\!\|$ instead of $\le \frac{1}{8}\|\!|x|\!\|$ on the left-hand side, which is possible by decomposing any $\psi \in \mathcal{P}^r$ as $\psi = \sum_{i=1}^N \psi_i^{1/4}$ where $\psi_i \in \mathcal{P}^r$ for some finite N depending only on d.)

For a distribution $\xi \in \mathcal{D}'(\Omega_b)$, define

$$\|\xi\|_{\mathring{\mathcal{C}}^\eta(\Omega_b)} = \sup_{x\in\Omega_b}\sup_{\psi\in\mathcal{P}^r}\sup_{\delta\in(0,\frac{1}{2}\|\!|x|\!\|)} \delta^{-\eta}|\langle \xi, \psi^\delta_x\rangle| = \sup_{x\in\Omega_b}\|\xi\|_{\eta;x;\le\frac{1}{2}\|\!|x|\!\|;r}\ .$$

Note that, for $\eta \le \theta$ and $h \in (0,1]$,

$$\|\xi\|_{\eta;x;\le h;r} \le h^{\theta-\eta}\|\xi\|_{\theta;x;\le h;r}\ .$$

Since $\eta \le \mathfrak{r} \le |\tau|_{\mathfrak{s}}$, it follows from (6.70) that, uniformly in $x \in \Omega_b$,

$$\|\tilde{\mathcal{R}}F\|_{\eta;x;\le\frac{1}{2}\|\!|x|\!\|;r} \lesssim \|\!|x|\!\|^{\omega-\eta+(\eta-\omega)} \max_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\tau,\omega,\eta}\|\Pi\|_{\tau;\Omega_b} + \max_{|\tau|_{\mathfrak{s}}<\omega} \|\!|x|\!\|^{|\tau|_s-\eta}|F^\tau_x|\,\|\Pi\|_{\tau;\Omega_b}\ ,$$

and thus

$$\|\tilde{\mathcal{R}}F\|_{\mathring{\mathcal{C}}^\eta(\Omega_b)} \lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} \|\Pi\|_{\tau;\Omega_b}\{|R^{\tau,F}|_{\tau,\omega,\eta} + \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\}\ .$$

The proof follows from Lemma A.2 (this is where we use $\eta \notin \mathbb{Z}$) which implies the existence of a bounded extension operator $\mathcal{E}\colon \mathring{\mathcal{C}}^\eta(\Omega_b) \to \mathcal{C}^\eta(\mathbb{R}^d)$ once we identify Ω_b with a shifted and dilated ball $\mathfrak{B}(0,b)$ and note that $\mathring{\mathcal{C}}^\eta$ embeds into $\widehat{\mathcal{C}}^\eta$ defined in Appendix A since $\lfloor 1-\eta\rfloor \ge r$. ∎

Remark 6.62. The extension $\mathcal{R}F = \mathcal{E}(\tilde{\mathcal{R}}F)$ is not unique in general.

6.8 Integration bounds

Fix in this subsection $0 < \lambda, b \le 1$, $\mathfrak{r} \in (-2, 0]$, $\eta \le \mathfrak{r}$ as in Lemma 6.61 and $\omega \in (0,1)$. Consider a model $Z = (\Pi, \Gamma)$ on Ω_b that is adapted to the rescaled kernel $K^\lambda(z) = \lambda^{|\mathfrak{s}|-2}K(\lambda \cdot_{\mathfrak{s}} z)$ from (6.35). Let $F \in \mathcal{D}^{\omega,\eta}(\Omega_b, \mathcal{V})$, where $\mathcal{V}$ is a sector of regularity $\mathfrak{r}$ on which the abstract integration map $\mathcal{I}$ is well-defined.[5] We define $\mathcal{K}^\lambda F \in \mathcal{D}^{\omega+2,\eta+2}_0(\Omega_b)$ as the integration operator from [Hai14, Prop. 6.16] or [GH19, Sec. 4.5] for the kernel K^λ and with input distribution $\mathcal{R}F \in \mathcal{C}^\eta(\mathbb{R}^d)$ from Lemma 6.61. Then, on Ω_b,

$$\tilde{\mathcal{R}}\mathcal{K}^\lambda F = K^\lambda * \mathcal{R}F\ .$$

The following is our main integration bound. The result is similar to others in the literature, e.g. [Hai14, GH19, BCZ24], but with the key remark that the bound is uniform in the parameter λ.

[5] $\mathcal{V}$ is allowed to contain polynomials by setting $\mathcal{I}[X^k] = 0$, in agreement with Definition 6.11. In later sections, we will always take $\mathcal{V} = \mathcal{W}_{\le 0}$.

Lemma 6.63. *With notation as above, denote*

$$\Theta = \max_{|\tau|_{\mathfrak{s}}<\omega} \|\Pi\|_{\tau;\lambda}\{|R^{\tau,F}|_{\tau,\omega,\eta} + \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\} \,,$$

where we recall $\|\Pi\|_{\tau;\lambda}$ *from* (6.40)*. Then, uniformly in* $\kappa \geq 0$ *and* $0 < \lambda, b \leq 1$*,*

$$|\mathcal{K}^\lambda F|_{\omega+2,\eta+2-\kappa} \lesssim b^\kappa \Theta \tag{6.71}$$

and for $l \in \mathbb{N}^d$ *with* $|l|_{\mathfrak{s}} < \omega + 2$*, uniformly in* $x \in \Omega_b$ *and* $0 < \lambda, b \leq 1$*,*

$$|\nabla^l (\mathcal{K}^\lambda F)_x| \lesssim (\|\!|x|\!\|^{\eta+2-|l|_{\mathfrak{s}}} + b^{\eta+2-|l|_{\mathfrak{s}}})\Theta \,. \tag{6.72}$$

Proof. The proof is a minor adaptation of [Hai14, Prop. 6.16] or [GH19, Lem. 4.12], so we only indicate the differences with [GH19, Lem. 4.12]. We restrict to $\omega < 1$ only because the polynomial sector of our regularity structure is spanned by X^l with $|l|_{\mathfrak{s}} \leq 2$.

To prove (6.71), it suffices to consider $\kappa = 0$ since $\|\!|x|\!\|^{\eta-\omega} \leq b^\kappa \|\!|x|\!\|^{\eta-\omega-\kappa}$ for all $\kappa \geq 0$ and $x \in \Omega_b$. Controlling $|\langle (\mathcal{K}^\lambda F)_y - \Gamma_{yz}(\mathcal{K}^\lambda F)_z, \sigma\rangle|$ for $|\sigma|_{\mathfrak{s}} \notin \mathbb{N}$ is simple and is done in the same way as in the proof of [Hai14, Thm. 5.12] (we use here Assumption 6.16). For $|\sigma|_{\mathfrak{s}} \in \mathbb{N}$, so that $\sigma = X^l$ for some $l \in \mathbb{N}^d$, inspecting the proof of [GH19, Lem. 4.12], we have uniformly in $x \in \Omega_b$

$$\begin{aligned}\sup_{y,z\in B_x(\frac{1}{2}\|\!|x|\!\|)} \sup_{l<\omega+2} \frac{|\langle (\mathcal{K}^\lambda F)_y - \Gamma_{yz}(\mathcal{K}^\lambda F)_z, X^l\rangle}{|y-z|_{\mathfrak{s}}^{\omega+2-l}} &\lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\omega-|\tau|_{\mathfrak{s}};B_x(\frac{1}{2}\|\!|x|\!\|)} \|\Pi\|_{\tau;\Omega_b} \\ &+ \max_{|\tau|_{\mathfrak{s}}<\omega} \|\!|x|\!\|^{|\tau|_{\mathfrak{s}}-\omega} \|F^\tau\|_{\infty;B_x(\frac{1}{2}\|\!|x|\!\|)} \|\Pi\|_{\tau;\lambda} \\ &+ \|\!|x|\!\|^{\eta-\omega} \|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)} \,.\end{aligned} \tag{6.73}$$

This bound arises from evaluating the various contributions of K_n^λ in the decomposition $K^\lambda = \sum_{n\geq 0} K_n^\lambda$ in the following way. Here we recall $K = \sum_{n\geq 0} K_n$ from (6.23) and denote $K_n^\lambda(z) = \lambda^{|\mathfrak{s}|-2} K_n(\lambda \cdot_{\mathfrak{s}} z)$, which is supported in $B_0(\lambda^{-1}2^{-n})$ and satisfies $\|D^k K_n^\lambda\|_\infty \lesssim (\lambda 2^n)^{|k|_{\mathfrak{s}}+|\mathfrak{s}|-2}$ for every $k \in \mathbb{N}^d$.

The first term on the right-hand side of (6.73) arises from those n such that $\lambda^{-1}2^{-n} \leq \frac{1}{2}\|\!|x|\!\|$, for which the situation is identical to the non-rescaled case and we can use the local reconstruction bound (6.69) and the argument in [GH19] or [Hai14].

The second term arises from bounding the contribution in [GH19, Eq. (4.38)] of $\Pi_y F_y$ evaluated on $D^{k+l}K_n^\lambda$ for $\lambda^{-1}2^{-n} > \frac{1}{2}\|\!|x|\!\|$. For this, we recall from (6.40) that $\|\Pi\|_{\tau;\lambda}$ bounds the evaluation of $\Pi_y\tau$ on non-anticipative test functions at scales up to λ^{-1}.

The third term arises in the same way as the second except we evaluate $\mathcal{R}F$ on $D^{k+l}K_n^\lambda$. For this, the contribution from n such that $\lambda^{-1}2^{-n} \leq 1$ is bounded by $\|\!|x|\!\|^{\eta-\omega}\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}$ in the same way as in [GH19] by exploiting that K_n^λ is essentially a non-anticipative test function at scale $\delta = \lambda^{-1}2^{-n}$ multiplied by δ^2. For $\lambda^{-1}2^{-n} > 1$, recall that $\mathcal{R}F$ is supported in $\operatorname{cl}\Omega_b$, therefore the contribution of each such term is of order $\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}(\lambda 2^n)^{|k+l|_{\mathfrak{s}}+|\mathfrak{s}|-2}$. Recall here that $|k+l|_{\mathfrak{s}} \geq \omega + 2$ and $\omega > 0$ and that this term has a prefactor $|y-z|_{\mathfrak{s}}^{|k+l|_{\mathfrak{s}}-\omega-2}$ in [GH19, Eq. (4.38)] with $|y-z| \lesssim \|\!|x|\!\| \leq b$. In particular, the exponent of $\lambda 2^n$ is positive and the prefactor is bounded above by 1, so the sum over $\lambda 2^n < 1$ gives a contribution of order $\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}$.

It follows from (6.73) that

$$\begin{aligned}|\mathcal{K}^\lambda F|_{\omega+2,\eta+2} &\lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} |R^{\tau,F}|_{\tau,\omega,\eta}\|\Pi\|_{\tau;\Omega_b} + \sup_{x\in\Omega_b} \max_{|\tau|_{\mathfrak{s}}<\omega} \|\!|x|\!\|^{|\tau|_{\mathfrak{s}}-\eta} |F_x^\tau|\, \|\Pi\|_{\tau;\lambda} + \|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)} \\ &\lesssim \max_{|\tau|_{\mathfrak{s}}<\omega} \|\Pi\|_{\tau;\lambda}\{|R^{\tau,F}|_{\tau,\omega,\eta} + \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\}\end{aligned}$$

where we used Lemma 6.61 and the fact that $\eta \le \mathfrak{r}$ in the final bound. This proves (6.71).

The bound (6.72) follows in a similar manner. Specifically, using that $\eta \le \mathfrak{r}$, one can check that the contribution to $|\langle \mathcal{K}^\lambda F(x), X^l\rangle$ for $|l|_{\mathfrak{s}} < 2+\omega$ of every term in the proof of [GH19, Lem. 4.12] is of order

$$|\!|\!| x |\!|\!|^{\eta+2-|l|_{\mathfrak{s}}}\Theta \tag{6.74}$$

except for $\mathcal{R}F(D^l K_n^\lambda(x-\cdot))$ appearing in [GH19, Eq. (4.36)] with $\lambda^{-1}2^{-n} > \frac{1}{2}|\!|\!| x |\!|\!|$. To bound the contributions of these final terms, we consider first the scales $\frac{1}{2}|\!|\!| x |\!|\!| < \lambda^{-1}2^{-n} \le 2b$, for which we use

$$|\mathcal{R}F(D^l K_n^\lambda(x-\cdot))| \lesssim (\lambda^{-1}2^{-n})^{\eta+2-|l|_{\mathfrak{s}}}\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}\,.$$

Since $\eta+2-|l|_{\mathfrak{s}} \notin \mathbb{Z}$ by assumption, summing over the relevant n gives a contribution of order

$$\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}(|\!|\!| x |\!|\!|^{\eta+2-|l|_{\mathfrak{s}}} + b^{\eta+2-|l|_{\mathfrak{s}}})\,. \tag{6.75}$$

For the scales $2b < \lambda^{-1}2^{-n}$, since $\mathcal{R}F$ is supported on $\operatorname{cl}\Omega_b$, it follows from Lemma 6.65 that

$$|\mathcal{R}F(D^l K_n^\lambda(x-\cdot))| \lesssim (\lambda^{-1}2^{-n})^{2-|l|_{\mathfrak{s}}-|\mathfrak{s}|}b^{\eta+|\mathfrak{s}|}\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}\,.$$

The exponent of $\lambda^{-1}2^{-n}$ is strictly negative since $|\mathfrak{s}| \ge 3$ (recall (1.7)), thus summing over n such that $\lambda^{-1}2^{-n} > 2b$ gives a contribution of order $b^{\eta+2-|l|_{\mathfrak{s}}}\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)}$. Since $\|\mathcal{R}F\|_{\mathcal{C}^\eta(\mathbb{R}^d)} \lesssim \Theta$ by Lemma 6.61, and combining with (6.74)-(6.75), the bound (6.72) follows. ∎

Remark 6.64. Bounding the contribution of $\mathcal{R}F$ on large scales is the only part where we require the extension of $\tilde{\mathcal{R}}F$. We suspect that a cleaner treatment would be to use the Dirichlet Green's function of $\mathcal{L}$ on Ω_b and leverage that it decays at the parabolic boundary, but this would require different arguments than in [Hai14, GH19].

Lemma 6.65. *Let $\eta \le 0$ and $\xi \in \mathcal{C}^\eta(\mathbb{R}^d)$ be supported on $\operatorname{cl}\mathfrak{B}(0,b)$ (recall (1.5)). Then, uniformly in $x \in \mathbb{R}^d$, $\psi \in \mathcal{B}^r$, and $\delta \ge b$,*

$$|\xi(\psi_x^\delta)| \lesssim b^{\eta+|\mathfrak{s}|}\delta^{-|\mathfrak{s}|}\|\xi\|_{\mathcal{C}^\eta(\mathbb{R}^d)}\,.$$

Proof. Since ξ has support in $\operatorname{cl}\mathfrak{B}(0,b)$, we can multiply ψ_x^δ with a suitable smooth cut-off function χ, which is 1 on $\mathfrak{B}(0,2b)$ and 0 outside $\mathfrak{B}(0,4b)$, so that $\chi\psi_x^\delta$ is a multiple of a test function at scale $4b$ multiplied by $b^{|\mathfrak{s}|}\delta^{-|\mathfrak{s}|}$ and $\xi(\chi\psi_x^\delta) = \xi(\psi_x^\delta)$. ∎

A useful consequence of the proof of Lemma 6.63 is the following

Lemma 6.66. *Let $\eta < 0$ with $\eta \notin \mathbb{Z}$ and $\xi\colon \Omega_b \to \mathbb{R}$ be measurable such that $\|\xi\|_{\infty;\eta} < \infty$ where we recall (6.67). Let $\mathcal{E}(\xi) \in \mathcal{C}^\eta(\mathbb{R}^d)$ be its extension from Lemma A.2. Then, uniformly in $x \in \Omega_b$ and $|l|_{\mathfrak{s}} \le 1$,*

$$|D^l K^\lambda * \mathcal{E}(\xi)(x)| \lesssim (|\!|\!| x |\!|\!|^{\eta+2-|l|_{\mathfrak{s}}} + b^{\eta+2-|l|_{\mathfrak{s}}})\|\xi\|_{\infty;\eta}\,.$$

Proof. Remark that $\|\mathcal{E}(\xi)\|_{\mathcal{C}^\eta(\mathbb{R}^d)} \lesssim \|\xi\|_{\mathring{\mathcal{C}}^\eta(\Omega_b)} \le \|\xi\|_{\infty;\eta}$, so indeed $\mathcal{E}(\xi)$ is well-defined.

We decompose as earlier $K^\lambda = \sum_{n\ge 0} K_n^\lambda$ where $\|D^l K_n^\lambda\|_\infty \lesssim (\lambda 2^n)^{|l|_{\mathfrak{s}}+|\mathfrak{s}|-2}$. Consider first the scales $\lambda^{-1}2^{-n} \le \frac{1}{2}|\!|\!| x |\!|\!|$ for which one has $|\langle D^l K_n^\lambda(x-\cdot), \xi\rangle| \lesssim |\!|\!| x |\!|\!|^\eta(\lambda 2^n)^{|l|_{\mathfrak{s}}-2}$. The exponent of $\lambda 2^n$ is strictly negative, so summing over the relevant n gives a contribution of order $|\!|\!| x |\!|\!|^{\eta-|l|_{\mathfrak{s}}+2}$. For the remaining scales $\lambda^{-1}2^{-n} > \frac{1}{2}|\!|\!| x |\!|\!|$, the contribution is of order $(|\!|\!| x |\!|\!|^{\eta+2-|l|_{\mathfrak{s}}} + b^{\eta+2-|l|_{\mathfrak{s}}})\|\mathcal{E}(\xi)\|_{\mathcal{C}^\eta(\mathbb{R}^d)}$ by the same argument as at the end of the proof of Lemma 6.63. ∎

6.9 Remainder bounds

We now make a choice for η. Let $\eta \in (-2,-1)$ such that, for all $\mathbf{k} \in B^\tau$,

$$\eta < \min \mathcal{A} = \min\{|\tau|_{\mathfrak{s}} \,:\, \tau \in \mathfrak{W}_{\le 0}\} \,, \tag{6.76}$$

$$\eta \le |\tau|_{\mathfrak{s}} - |\mathbf{k}|_{\mathfrak{s}} - \nabla\eta(\tau,\mathbf{k}) \,, \tag{6.77}$$

where we recall $\nabla\eta(\tau,\mathbf{k}) \ge 0$ from Assumption 6.22 (such η exists by (6.14) in Assumption 6.22). Note that $\lfloor 1-\eta \rfloor = 2 \ge r$ as required in Lemma 6.61.

In this subsection, we fix $0 < \lambda, b \le 1$ and a model (Π,Γ) on Ω_b adapted to K^λ. Define

$$\|(f,\Gamma)\| = \max_{\mathcal{I}^k\sigma \in \mathfrak{U}} |\!|\!|\Upsilon^\sigma|\!|\!| \|\Gamma\|_{\mathcal{I}^k\sigma;\Omega_b} \,.$$

Let φ be a solution to the remainder of (6.1) in the sense of Definition 6.32 such that $\|\varphi\|_\infty \le \varepsilon$. Define $F\colon \Omega_b \to \mathcal{W}_{\le 0} \otimes E$ by

$$F_x = \sum_{\tau \in \mathfrak{W}_{\le 0}} F_x^\tau \tau \,, \quad \text{where } \; F_x^\tau = \frac{1}{\tau!}\Upsilon^\tau(\boldsymbol{\varphi}_x) \,,$$

so that $\langle F_x, \tau\rangle = \Upsilon^\tau(\boldsymbol{\varphi}_x)$, and denote

$$R_{x,y}^\tau = R_{x,y}^{\tau,F} = \frac{1}{\tau!}\langle F_y - \Gamma_{yx}F_x, \tau\rangle \,,$$

so that $\sum_{\tau\in\mathfrak{W}_{\le 0}} R_{x,y}^\tau \tau = F_y - \Gamma_{yx}F_x$. Define

$$V = \max_{\tau\in\mathfrak{W}_{\le 0}} |R^\tau|_{\tau,\zeta,\eta} \|\Pi\|_{\tau;\lambda} \,, \qquad W = \max_{\tau\in\mathfrak{W}_{\le 0}} \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}} \|\Pi\|_{\tau;\lambda} \,,$$

where we recall ζ from (6.10).

Remark 6.67. For all $\tau \in \mathfrak{W}_{<0}$, recalling Φ from (6.25), one has

$$\langle F_x - \Gamma_{xy}F_y, \tau\rangle = \langle \Phi_x - \Gamma_{xy}\Phi_y, \mathcal{I}[\tau]\rangle \,.$$

Moreover, φ and $\nabla\varphi$ are ζ-Hölder continuous since $\Phi \in \mathcal{D}^{2+\zeta}(\Omega_b, E)$ by assumption. Hence

$$|R_{x,y}^{\mathbf{1}}| = |P(\boldsymbol{\varphi}_y) - P(\boldsymbol{\varphi}_x)| \lesssim |x-y|_{\mathfrak{s}}^\zeta \,.$$

Therefore $F \in \mathcal{D}^\zeta(\Omega_b, E)$ and $V, W < \infty$.

Recall the set of multi-indexes B^τ from Definition 6.20.

Lemma 6.68. *Suppose $\varepsilon, V, W \le 1$ and $b \le 1/8$. Define*

$$\Lambda = \varepsilon + b^{\eta+2}(V + W + \|(f,\Gamma)\|) \,.$$

Recall the quantities $|\!|\!|D^{\mathbf{k}}\Upsilon^\tau|\!|\!|$ from Assumption 6.22. Then, for $\tau \in \mathfrak{W}_{\le 0}$,

$$\|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}} \lesssim |\!|\!|\Upsilon^\tau|\!|\!| \,, \qquad \|F^{\mathbf{1}}\|_{\infty;\eta} \le \|F^{\mathbf{1}}\|_{\infty;-1} \lesssim \varepsilon(\varepsilon + b^{\eta+2}(V+W)) \tag{6.78}$$

and

$$|R^\tau|_{\tau,\zeta,\eta} \lesssim \max_{\mathbf{k}\in B^\tau\,:\,\mathbf{k}\ne 0} |\!|\!|D^{\mathbf{k}}\Upsilon^\tau|\!|\!| \Lambda^{|\mathbf{k}|} \,. \tag{6.79}$$

The proportionality constants above do not depend on $b, \lambda, \varepsilon, f, Z, \varphi$.

Proof. By Definition 6.32 and (6.69), we have $\mathcal{L}\varphi = \tilde{\mathcal{R}}F$ in Ω_b. Moreover, $\mathcal{L}K^\lambda = \delta_0 + U^\lambda$ where U^λ is smooth, compactly supported, and vanishes on $B_0(\lambda^{-1}/2)$. Define $Y\colon \Omega_b \to E$ by

$$\varphi = K^\lambda * \mathcal{R}F + Y \,. \tag{6.80}$$

Then, on Ω_b,

$$\mathcal{L}Y + U^\lambda * \mathcal{R}F = 0 \,.$$

Since $\mathcal{R}F$ is supported on $\operatorname{cl}\Omega_b$, which has diameter $b \le 1/8$, and U^λ is compactly supported outside $B_0(1/2)$, we have $U^\lambda * (\mathcal{R}F) \equiv 0$ on Ω_b and thus $\mathcal{L}Y = 0$.

Our conditions on η (including (6.76)) imply that we are in the scope of the integration Lemma 6.63. Therefore, by the integration bound (6.72) with $l = 0$, since $\eta + 2 > 0$,

$$\|K^\lambda * \mathcal{R}F\|_{\infty;\Omega_b} \lesssim b^{\eta+2} \max_{|\tau|_{\mathfrak{s}}<\zeta} \|\Pi\|_{\tau;\lambda}\{|R^\tau|_{\tau,\zeta,\eta} + \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\} \le b^{\eta+2}(V+W) \,.$$

Therefore, since $\|\varphi\|_{\infty;\Omega_b} \le \varepsilon$ by assumption,

$$\|Y\|_{\infty;\Omega_b} \lesssim \varepsilon + b^{\eta+2}(V+W) \,.$$

Therefore, since $\mathcal{L}Y = 0$ in Ω_b, it follows from classical weighted Schauder estimates [Lie96, Thm. 4.9, p.59] that, for any $|k|_{\mathfrak{s}} \le 2$ and $\kappa \in [0,1)$, uniformly in $x \in \Omega_b$,

$$\|D^k Y\|_{\mathcal{C}^\kappa; B_x(\frac{1}{2}|||x|||)} \lesssim |||x|||^{-|k|_{\mathfrak{s}}-\kappa}(\varepsilon + b^{\eta+2}(V+W)) \,. \tag{6.81}$$

Moreover, one has the equality between modelled distributions in $\mathcal{D}_0^{\zeta+2,\eta+2}(\Omega_b)$

$$\Phi = \mathcal{K}^\lambda F + Y \,,$$

where we recall Φ from (6.25) and lift Y canonically to the polynomial regularity structure.

Next, for $|k|_{\mathfrak{s}} \le 2$, by the integration bound (6.72) and by (6.81), we have

$$|(\nabla^k \varphi)_x| \lesssim |||x|||^{-|k|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W)) \lesssim |||x|||^{-|k|_{\mathfrak{s}}} \,, \tag{6.82}$$

where the final bound follows from our assumption that $\varepsilon, V, W, b \le 1$. It follows that

$$|D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi}_x)| \lesssim |||D^{\mathbf{k}}\Upsilon^\tau|||(1 + |\nabla\varphi_x|^{\nabla\eta(\tau,\mathbf{k})}) \lesssim |||D^{\mathbf{k}}\Upsilon^\tau|||\,|||x|||^{-\nabla\eta(\tau,\mathbf{k})} \lesssim |||D^{\mathbf{k}}\Upsilon^\tau|||\,|||x|||^{\eta-|\tau|_{\mathfrak{s}}+|\mathbf{k}|_{\mathfrak{s}}} \tag{6.83}$$

where we used (6.11) in Assumption 6.22 and that $\eta(\tau,\mathbf{k}), \bar{\eta}(\tau,\mathbf{k}) \ge 0$ for the first bound, (6.82) and that $\nabla\eta(\tau,\mathbf{k}) \ge 0$ for the second bound, and the condition (6.77) on η for the third bound.

In particular,

$$|\Upsilon^\tau(\boldsymbol{\varphi}_x)| \lesssim |||\Upsilon^\tau|||\,|||x|||^{\eta-|\tau|_{\mathfrak{s}}} \lesssim |||\Upsilon^\tau|||b^{\eta+2}|||x|||^{-|\tau|_{\mathfrak{s}}-2} \,. \tag{6.84}$$

Moreover, for $\tau = \mathbf{1}$, so that $\Upsilon^{\mathbf{1}} = P$, using Assumption 6.1 on P and that $|\varphi_x| \le \varepsilon \le 1$, this can be improved to

$$|\Upsilon^{\mathbf{1}}(\boldsymbol{\varphi}_x)| \lesssim |\varphi_x|^2 + |\varphi_x|\,|\nabla\varphi_x| \lesssim |||x|||^{-1}\varepsilon(\varepsilon + b^{\eta+2}(V+W)) \,. \tag{6.85}$$

We obtain (6.78) from (6.85) and the first bound in (6.84).

It remains to prove (6.79). Fix $\tau \in \mathfrak{W}_{\le 0}$. Recall $\mathring{B}^\tau$ from Definition 6.20. Define further

$$\partial\mathring{B}^\tau = \{\mathbf{k} \in \mathbb{N}^{\mathbb{N}^d_{<2}} : \mathbf{k} \notin \mathring{B}^\tau \text{ and } \mathbf{k} \setminus \{k\} \in \mathring{B}^\tau \text{ for all } k \in \mathbf{k} \text{ and } D^{\mathbf{k}}\Upsilon^\tau \neq 0\} \,,$$

so that $B^\tau = \mathring{B}^\tau \sqcup \partial\mathring{B}^\tau$. By the generalised Taylor formula [Hai14, Prop. 11.1],[6]

$$\Upsilon^\tau(\boldsymbol{\varphi}_y) = \sum_{\mathbf{k}\in\mathring{B}^\tau} \frac{D^{\mathbf{k}}\Upsilon^\tau}{\mathbf{k}!}(\boldsymbol{\varphi}_x)(\boldsymbol{\varphi}_y - \boldsymbol{\varphi}_x)^{\mathbf{k}} + O\big(\max_{\mathbf{k}\in\partial\mathring{B}^\tau} \|D^{\mathbf{k}}\Upsilon^\tau\|_{\infty;B_x(\frac{1}{2}|||x|||)}(\boldsymbol{\varphi}_y - \boldsymbol{\varphi}_x)^{\mathbf{k}}\big)\,, \tag{6.86}$$

where we denote

$$\|D^{\mathbf{k}}\Upsilon^\tau\|_{\infty;B_x(\frac{1}{2}|||x|||)} = \sup_{z\in B_x(\frac{1}{2}|||x|||)} |D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi}_z)|$$

and for $\mathbf{k} = \{k_1,\dots,k_{|\mathbf{k}|}\} \in \mathbb{N}^{\mathbb{N}^d_{<2}}$ and $\boldsymbol{\varphi} = \sum_{|l|_{\mathfrak{s}}\le 2} \frac{1}{l!}(\nabla^l\varphi)X^l \in \bar{\mathcal{T}}_{\le 2}\otimes E$, we write $\boldsymbol{\varphi}^{\mathbf{k}} = (\nabla^{k_1}\varphi,\dots,\nabla^{k_{|\mathbf{k}|}}\varphi)$, which we treat as a tensor in $E^{\otimes|\mathbf{k}|}$.

For $|k|_{\mathfrak{s}} < 2$, define the remainder

$$\nabla^k Q_{x,y} = \langle \mathscr{D}^k\boldsymbol{\varphi}_y - \Gamma_{yx}\mathscr{D}^k\boldsymbol{\varphi}_x, \mathbf{1}\rangle$$

where

$$\mathscr{D}^k \colon \mathcal{I}[\mathcal{W}_{<0}]\oplus\bar{\mathcal{T}}_{\le 2} \to \mathcal{I}^k[\mathcal{W}_{<0}]\oplus\bar{\mathcal{T}}_{\le 2-|k|_{\mathfrak{s}}}$$

is the linear map for which $\mathscr{D}^k\mathcal{I}[\tau] = \mathcal{I}^k[\tau]$ and $\mathscr{D}^k X^m = \frac{m!}{(m-k)!}X^{m-k}$ if $m_i \ge k_i$ for all $i\in[d]$ and $\mathscr{D}^k X^m = 0$ otherwise. Then

$$\nabla^k Q_{x,y} = \nabla^k\varphi_y - \nabla^k\varphi_x - \sum_{\mathcal{I}^k[\sigma]\in\mathfrak{U}} \frac{1}{\sigma!}\Upsilon^\sigma(\boldsymbol{\varphi}_x)\langle\Gamma_{yx}\mathcal{I}^k[\sigma],\mathbf{1}\rangle - \sum_{0<|l|_{\mathfrak{s}}\le\zeta+2-|k|_{\mathfrak{s}}} \frac{1}{l!}(\nabla^{k+l}\varphi)_x(y-x)^l\,. \tag{6.87}$$

Recall that we fixed $\tau\in\mathfrak{W}_{\le 0}$. Then

$$R^\tau_{x,y} = \frac{1}{\tau!}\langle F_y - \Gamma_{yx}F_x, \tau\rangle = \frac{1}{\tau!}\Big(\Upsilon^\tau(\boldsymbol{\varphi}_y) - \sum_{\sigma\in\mathfrak{W}_{\le 0}} \frac{1}{\sigma!}\Upsilon^\sigma(\boldsymbol{\varphi}_x)\langle\tau,\Gamma_{yx}\sigma\rangle\Big)\,.$$

Recall that (6.54) in Lemma 6.53 gives another expression for the final sum. Substituting into (6.86) the expression for $\nabla^k\varphi_y - \nabla^k\varphi_x$ that we obtain by rearranging (6.87), we see that the terms arising from (6.54) cancel the lowest order terms of $\Upsilon^\tau(\boldsymbol{\varphi}_y)$ in (6.86) (for this, recall by Remark 6.54 that $\mathbf{k}+\mathbf{m}\in\mathring{B}^\tau$ for all $\mathbf{k},\mathbf{m}$ appearing in (6.54)). Therefore, for $y\in B_x(\frac{1}{2}|||x|||)$, $R^\tau_{x,y}$ is made up of a term of order

$$\max_{\mathbf{k}\in\partial\mathring{B}^\tau} \|D^{\mathbf{k}}\Upsilon^\tau\|_{\infty;B_x(\frac{1}{2}|||x|||)}|(\boldsymbol{\varphi}_y - \boldsymbol{\varphi}_x)^{\mathbf{k}}| \tag{6.88}$$

plus a finite sum of terms which are multiples of

$$D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi}_x)\Big(\prod_{1\le i\le q}\nabla^{k_i}Q_{x,y}\Big)\Big(\prod_{q<i\le j}\Upsilon^{\tau_i}(\boldsymbol{\varphi}_x)\langle\Gamma_{yx}\mathcal{I}^{k_i}\tau_i,\mathbf{1}\rangle\Big)\Big(\prod_{j<i\le|\mathbf{k}|}(\nabla^{k_i+l_i}\varphi)_x(y-x)^{|l_i|_{\mathfrak{s}}}\Big) \tag{6.89}$$

where $\mathbf{k} = \{k_1,\dots,k_{|\mathbf{k}|}\}\in\mathring{B}^\tau$ with $\mathbf{k}\ne 0$, $0\le q\le j\le|\mathbf{k}|$, $\mathcal{I}^{k_i}\tau_i\in\mathfrak{U}$, and $l_i\in\mathbb{N}^d$ subject to $0<|l_i|_{\mathfrak{s}}\le 2-|k_i|_{\mathfrak{s}}$ and the following conditions:

- $q\ge 1$, or
- $q=0$ and $|\tau|_{\mathfrak{s}} + \sum_{0<i\le j}|\mathcal{I}^{k_i}\tau_i|_{\mathfrak{s}} + \sum_{j<i\le|\mathbf{k}|}|l_i|_{\mathfrak{s}} > 0$.

[6] To apply [Hai14, Prop. 11.1], we put an arbitrary total order $<$ on $\mathbb{N}^d_{<2}$ such that $k<\bar{k}$ implies $\gamma_k\le\gamma_{\bar{k}}$, so that our definition of $\partial\mathring{B}^\tau$ coincides with that of [Hai14, Appendix A] modulo multi-indexes $\mathbf{k}\in\mathbb{N}^{\mathbb{N}^d_{<2}}$ such that $D^{\mathbf{k}}\Upsilon^\tau = 0$, which can be ignored for our argument.

We proceed to bound (6.88)-(6.89).

First, for $|k|_{\mathfrak{s}} < 2$, by the fact that $\mathscr{D}^k$ commutes with Γ_{yx}, by the integration bound (6.71) with $\kappa = \eta + 2$, and by (6.81), we obtain for $y \in B_x(\frac{1}{2}\|x\|)$

$$|\nabla^k Q_{x,y}| \lesssim \|x\|^{-\zeta-2}|x-y|_{\mathfrak{s}}^{\zeta+2-|k|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W)) \,. \tag{6.90}$$

Then, by (6.87), (6.90), (6.84), and (6.82), for $|k|_{\mathfrak{s}} < 2$,

$$\begin{aligned}
|\nabla^k \varphi_y - \nabla^k \varphi_x| &\lesssim \|x\|^{-\zeta-2}|x-y|^{\zeta+2-|k|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W)) \\
&\quad + \max_{\mathcal{I}^k[\sigma]\in\mathfrak{U}} b^{\eta+2}\|x\|^{-|\sigma|_{\mathfrak{s}}-2}|x-y|_{\mathfrak{s}}^{|\sigma|_{\mathfrak{s}}+2-|k|_{\mathfrak{s}}}\|\Upsilon^\sigma\|\|\Gamma\|_{\mathcal{I}^k[\sigma]} \\
&\quad + \max_{0<|l|_{\mathfrak{s}}<\zeta+2-|k|_{\mathfrak{s}}} \|x\|^{-|k|_{\mathfrak{s}}-|l|_{\mathfrak{s}}}|x-y|^{|l|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W)) \\
&\lesssim \big\{\varepsilon + b^{\eta+2}(V + W + \|(f,\Gamma)\|)\big\}\|x\|^{-|k|_{\mathfrak{s}}-\gamma_k}|x-y|_{\mathfrak{s}}^{\gamma_k} \\
&= \Lambda\|x\|^{-|k|_{\mathfrak{s}}-\gamma_k}|x-y|_{\mathfrak{s}}^{\gamma_k}
\end{aligned} \tag{6.91}$$

where in the second bound we used that $\zeta + 2 - |k|_{\mathfrak{s}}, |l|_{\mathfrak{s}} \geq 1 \geq \gamma_k$ and that $|\sigma|_{\mathfrak{s}} + 2 - |k|_{\mathfrak{s}} = |\mathcal{I}^k[\sigma]|_{\mathfrak{s}} > 0$ implies $|\mathcal{I}^k[\sigma]|_{\mathfrak{s}} \geq \gamma_k$ by definition of γ_k from (6.9).

We can now bound the first remainder term (6.88). By (6.83) and (6.91), for $\mathbf{k} \in \partial\mathring{B}^\tau$,

$$\begin{aligned}
\|D^{\mathbf{k}}\Upsilon^\tau\|_{\infty;B_x(\frac{1}{2}\|x\|)}|(\boldsymbol{\varphi}_y - \boldsymbol{\varphi}_x)^{\mathbf{k}}| &\lesssim \|D^{\mathbf{k}}\Upsilon^\tau\|\|x\|^{\eta-|\tau|_{\mathfrak{s}}+|\mathbf{k}|_{\mathfrak{s}}}|(\boldsymbol{\varphi}_y - \boldsymbol{\varphi}_x)^{\mathbf{k}}| \\
&\lesssim \|D^{\mathbf{k}}\Upsilon^\tau\|\Lambda^{|\mathbf{k}|}\|x\|^{\eta-|\tau|_{\mathfrak{s}}-\gamma_{\mathbf{k}}}|x-y|_{\mathfrak{s}}^{\gamma_{\mathbf{k}}} \\
&\leq \|D^{\mathbf{k}}\Upsilon^\tau\|\Lambda^{|\mathbf{k}|}\|x\|^{\eta-\zeta}|x-y|_{\mathfrak{s}}^{\zeta-|\tau|_{\mathfrak{s}}}
\end{aligned}$$

where we used that $\gamma_{\mathbf{k}} \geq \zeta - |\tau|_{\mathfrak{s}}$ by definition of $\partial\mathring{B}^\tau$.

Turning to the terms (6.89), suppose first that $q \geq 1$. Then, by (6.90), (6.84), and (6.82),

$$\begin{aligned}
(6.89) &\lesssim |D^{\mathbf{k}}\Upsilon^\tau(\boldsymbol{\varphi}_x)|\Big\{\prod_{1\leq i\leq q}\|x\|^{-\zeta-2}|x-y|_{\mathfrak{s}}^{\zeta+2-|k_i|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W))\Big\} \\
&\quad\times\Big\{\prod_{q<i\leq j}\|\Upsilon^{\tau_i}\|b^{\eta+2}\|x\|^{-|\tau_i|_{\mathfrak{s}}-2}\|\Gamma\|_{\mathcal{I}^{k_i}[\tau_i]}|x-y|_{\mathfrak{s}}^{|\tau_i|_{\mathfrak{s}}+2-|k_i|_{\mathfrak{s}}}\Big\} \\
&\quad\times\Big\{\prod_{j<i\leq|\mathbf{k}|}\|x\|^{-|k_i|_{\mathfrak{s}}-|l_i|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W))|y-x|_{\mathfrak{s}}^{|l_i|_{\mathfrak{s}}}\Big\} \,.
\end{aligned} \tag{6.92}$$

By Lemma 6.58, since $D^{\mathbf{k}}\Upsilon^\tau \neq 0$, one has $|\tau|_{\mathfrak{s}} - \beta + \sum_{k\in\mathbf{k}}\{\beta + 2 - |k_i|_{\mathfrak{s}}\} \geq 0$, so in particular $2 - |k_1|_{\mathfrak{s}} \geq -|\tau|_{\mathfrak{s}}$ (recall that $\mathbf{k} \neq 0$). Furthermore, the exponents of $\|x\|$ and $|x-y|_{\mathfrak{s}}$ in the three parentheses in (6.92) sum to $-|\mathbf{k}|_{\mathfrak{s}}$ and, since $q \geq 1$, the total exponent of $|x-y|_{\mathfrak{s}}$ is at least $\zeta + 2 - |k_1|_{\mathfrak{s}} \geq \zeta - |\tau|_s$. Since $|x-y|_{\mathfrak{s}} \leq \|x\|$, it follows from (6.83) that

$$\begin{aligned}
(6.89) &\lesssim \|D^{\mathbf{k}}\Upsilon^\tau\|\|x\|^{\eta-\zeta}|x-y|_{\mathfrak{s}}^{\zeta-|\tau|_{\mathfrak{s}}}(\varepsilon + b^{\eta+2}(V+W))^{q+|\mathbf{k}|-j}(b^{\eta+2}\|(f,\Gamma)\|)^{j-q} \\
&\leq \|x\|^{\eta-\zeta}|x-y|_{\mathfrak{s}}^{\zeta-|\tau|_{\mathfrak{s}}}\|D^{\mathbf{k}}\Upsilon^\tau\|\Lambda^{|\mathbf{k}|} \,.
\end{aligned} \tag{6.93}$$

Suppose now that $q = 0$. Since $D^{\mathbf{k}}\Upsilon^\tau \neq 0$, by the same reasoning as in the proof of Lemma 6.58, there exists a conforming tree $\sigma \in \mathfrak{W}$ formed by grafting the family of trees $\{\mathcal{I}^{k_i}[\tau_i]\}_{0<i\leq j}$ onto vertices of τ, and for which $|\sigma|_{\mathfrak{s}} = |\tau|_{\mathfrak{s}} + \sum_{0<i\leq j}|\mathcal{I}^{k_i}[\tau_i]|_{\mathfrak{s}}$. Therefore, by our choice of $\zeta \in (0,1)$ in (6.10) and the facts that $|l_i|_{\mathfrak{s}} \in \mathbb{N}\setminus\{0\}$ and $|\mathbf{k}| \neq 0$,

$$|\tau|_{\mathfrak{s}} + \sum_{0<i\leq j}|\mathcal{I}^{k_i}[\tau_i]|_{\mathfrak{s}} + \sum_{j<i\leq|\mathbf{k}|}|l_i|_{\mathfrak{s}} > 0 \quad\text{implies}\quad |\tau|_{\mathfrak{s}} + \sum_{0<i\leq j}|\mathcal{I}^{k_i}[\tau_i]|_{\mathfrak{s}} + \sum_{j<i\leq|\mathbf{k}|}|l_i|_{\mathfrak{s}} \geq \zeta \,.$$

Then we obtain the same bound (6.93) in a similar way as for the case $q \geq 1$.

In conclusion, $|R^\tau|_{\tau,\zeta,\eta} \lesssim \max_{\mathbf{k}\in B^\tau\,:\,\mathbf{k}\neq 0}\|D^{\mathbf{k}}\Upsilon^\tau\|\Lambda^{|\mathbf{k}|}$, which is precisely (6.79). ∎

Define now

$$V^* = \max_{\tau\in\mathfrak{W}_{<0}} |R^\tau|_{\tau,\zeta,\eta}\|\Pi\|_{\tau;\lambda}\,, \qquad W^* = \max_{\tau\in\mathfrak{W}_{<0}} \|F^\tau\|_{\infty;\eta-|\tau|_{\mathfrak{s}}}\|\Pi\|_{\tau;\lambda}\,,$$

$$\|(f,\Pi)\| = \max_{|\tau|_{\mathfrak{s}}<0} \|\Pi\|_{\tau;\lambda}|\!|\!|\Upsilon^\tau|\!|\!|\,, \qquad |\!|\!|(f,\Pi)|\!|\!| = \max_{|\tau|_{\mathfrak{s}}<0}\max_{\mathbf{k}\in B^\tau:\mathbf{k}\neq 0} \|\Pi\|_{\tau;\lambda}|\!|\!|D^{\mathbf{k}}\Upsilon^\tau|\!|\!|\,,$$

$$\|(f,Z)\| = \|(f,\Gamma)\| + \|(f,\Pi)\| + |\!|\!|(f,\Pi)|\!|\!|\,.$$

Lemma 6.69. *There exists $\tilde\delta > 0$ such that, if $\|(f,Z)\|, b, \varepsilon < \tilde\delta$, then*

$$V + W \le 1/2\,, \tag{6.94}$$

$$V^* \le |\!|\!|(f,\Pi)|\!|\!|\,, \qquad W^* \lesssim \|(f,\Pi)\|\,, \tag{6.95}$$

where the proportionality constants do not depend on $b, \lambda, \varepsilon, f, Z, \varphi$.

Proof. Suppose $b, \varepsilon, \|(f,\Gamma)\|, |\!|\!|(f,\Pi)|\!|\!| \le \tilde\delta \ll 1$. Suppose also first that $V + W \le 1$. Recall $\Lambda = \varepsilon + b^{\eta+2}(V + W + \|(f,\Gamma)\|)$. Multiplying (6.79) by $\|\Pi\|_{\tau;\lambda}$ and taking the max over τ, we obtain by Lemma 6.68 that, for $\tilde\delta > 0$ sufficiently small,

$$V \le C \max_{\tau\in\mathfrak{W}_{\le 0}}\max_{\mathbf{k}\in B^\tau:\mathbf{k}\neq 0} \|\Pi\|_{\tau;\lambda}|\!|\!|D^{\mathbf{k}}\Upsilon^\tau|\!|\!|\Lambda^{|\mathbf{k}|} \le 1/4\,, \tag{6.96}$$

where $C > 0$ does not depend on $b, Z, f, \lambda, \varepsilon, \varphi$. If in addition $\|(f,\Pi)\| \le \tilde\delta$, then, by (6.78),

$$W \le C\varepsilon(\varepsilon + b^{\eta+2}(V+W)) + C\|(f,\Pi)\| \le 1/4\,.$$

This proves (6.94) under the assumption that $V + W \le 1$.

On the other hand, remark that if φ is a solution for $b > 0$, then it is also a solution for every $\bar b < b$. Let V^b, W^b be the corresponding quantities as functions of b. Then V^b, W^b are continuous and increasing in b since the coefficients $\boldsymbol{\varphi}_x$ are continuous in x (to see this for V^b, we make use of the weight $|\!|\!|x|\!|\!|^{\zeta-\eta}$ in the definition (6.68) of $|R|_{\tau,\zeta,\eta}$). Moreover, since $\eta - |\tau|_{\mathfrak{s}} < 0$ for every $\tau \in \mathfrak{W}_{\le 0}$ due to condition (6.76), we have $V^b, W^b \to 0$ as $b \to 0$. Therefore, for all b sufficiently small, possibly depending on φ, we can ensure that $V^b + W^b \le 1/2$. We thus obtain (6.94) by continuity in b without the prior assumption that $V + W \le 1$.

The bound on W^* follows directly from the first bound in (6.78). To bound V^*, note that the first inequality in (6.96) holds with V replaced by V^* and where the first maximum is taken over $\tau \in \mathfrak{W}_{<0}$. The bound on V^* then follows by taking $\tilde\delta$ small so that $\Lambda \ll 1$. ∎

6.10 Local coercivity

We now verify the local coercivity condition of Corollary 3.5, which concludes the proof of Theorem 6.35. Recall the notation from Section 6.4.

Lemma 6.70 (Local coercivity)**.** *There exists $b, r, \varepsilon, \delta > 0$ sufficiently small such that, if $M = (f,Z) \in \mathbb{M}_{z,\lambda}$ with $\|M\|_{z,\lambda} \le r$ and $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_{\infty;\Omega_b} \le \varepsilon$, then $|\varphi(0)| < \varepsilon - \delta$.*

Proof. We use notation from Sections 6.7-6.9. Let $\tilde\delta > 0$ be as in Lemma 6.69. We take $b, \varepsilon, r \le \tilde\delta$ sufficiently small and $M = (f,Z) \in \mathbb{M}_{z,\lambda}$ such that $\|M\|_{z,\lambda} \le r$. Let $\varphi \in \mathbb{S}_M$ with $\|\varphi\|_{\infty;\Omega_b} \le \varepsilon$. Below, all implicit proportionality constants do not depend on $b, \varepsilon, r, \lambda, M, \varphi$.

By Assumption 6.4 and a scaling argument, there exist $\delta > 0$ and $\mathring\varphi \in \mathcal{C}(\mathrm{cl}\,\Omega_b, E)$ such that $\mathring\varphi \restriction_{\partial\Omega_b} = \varphi \restriction_{\partial\Omega_b}$, $\mathcal{L}\mathring\varphi = P(\mathring\varphi)$ on Ω_b, and $|\mathring\varphi(0)| < \varepsilon - 2\delta$. Our goal is to estimate $\varphi - \mathring\varphi$.

We lift $\mathring{\varphi}$ canonically to a modelled distribution $\mathring{\boldsymbol{\varphi}}\colon \Omega_b \to \bar{\mathcal{T}}_{\leq 2} \otimes E$ taking values in the polynomial regularity structure with $\langle \mathring{\boldsymbol{\varphi}}_x, X^k \rangle = (\nabla^k \mathring{\varphi})_x$. Denote

$$\|\mathring{\boldsymbol{\varphi}}\|_{\infty;<2} = \max_{|k|_{\mathfrak{s}}<2} \sup_{x\in\Omega_b} |\!|\!| x |\!|\!|^{|k|_{\mathfrak{s}}} |\nabla^k \mathring{\varphi}_x| \, .$$

Then, since $\|\mathring{\varphi}\|_{\infty;\partial\Omega_b} \leq \varepsilon$ and by stability of the equation $\mathcal{L}\mathring{\varphi} = P(\mathring{\varphi})$ at the zero solution,

$$\|\mathring{\boldsymbol{\varphi}}\|_{\infty;<2} \lesssim \varepsilon \, . \tag{6.97}$$

As in (6.80) define $Y, \mathring{Y} \in \mathcal{C}(\mathrm{cl}\,\Omega_b, E)$ by

$$\varphi = K^\lambda * \mathcal{R}\{P(\boldsymbol{\varphi})\mathbf{1} + \mathcal{R}\tilde{F}\} + Y \, , \quad \mathring{\varphi} = K^\lambda * \mathcal{R}\{P(\mathring{\boldsymbol{\varphi}})\mathbf{1}\} + \mathring{Y} \tag{6.98}$$

where we write $\tilde{F} = \sum_{|\tau|_{\mathfrak{s}}<0} F^\tau \tau$ and where $\mathcal{R}$ is the same reconstruction operator as in (6.80), i.e. the operator from Lemma 6.61 for the domain Ω_b and for $\eta \in (-2,-1)$ from Section 6.9.

NB. We apply Lemma 6.61 with $\eta \in (-2,-1)$, therefore $\mathcal{R}\{P(\mathring{\boldsymbol{\varphi}})\mathbf{1}\}$ is in $\mathcal{C}^\eta(\mathbb{R}^d)$ but not necessarily in $\mathcal{C}^\omega(\mathbb{R}^d)$ for any $\omega \geq -1$.

By the same argument as at the start of the proof of Lemma 6.68, on Ω_b

$$\mathcal{L}Y = \mathcal{L}\mathring{Y} = 0 \, .$$

We furthermore have the equality between modelled distributions on Ω_b

$$\Phi = \mathcal{K}^\lambda(P(\boldsymbol{\varphi})\mathbf{1} + \tilde{F}) + Y \, , \qquad \mathring{\boldsymbol{\varphi}} = \mathcal{K}^\lambda(P(\mathring{\boldsymbol{\varphi}})\mathbf{1}) + \mathring{Y} \, ,$$

where we lift $Y, \mathring{Y}$ canonically to the polynomial regularity structure and $\mathcal{K}^\lambda$ is the integration map as in Section 6.8 that uses $\mathcal{R}$ as input.

In a similar way as in (6.85), using (6.82) and (6.97), we have

$$\begin{aligned} |P(\boldsymbol{\varphi}_x) - P(\mathring{\boldsymbol{\varphi}}_x)| &\lesssim |\!|\!| x |\!|\!|^{-1} \|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} (\|\boldsymbol{\varphi}\|_{\infty;<2} + \|\mathring{\boldsymbol{\varphi}}\|_{\infty;<2}) \\ &\lesssim |\!|\!| x |\!|\!|^{-1} \|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} (\varepsilon + b^{\eta+2}(V+W)) \\ &\lesssim |\!|\!| x |\!|\!|^{\eta} \|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} (\varepsilon + b^{\eta+2}(V+W)) \, , \end{aligned}$$

where we used $\eta < -1$ in the final bound. Since $\eta > -2$, we obtain by Lemma 6.66

$$\|K^\lambda * \mathcal{R}(P(\boldsymbol{\varphi})\mathbf{1} - P(\mathring{\boldsymbol{\varphi}})\mathbf{1})\|_{\infty;<2} \lesssim b^{\eta+2} \|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} (\varepsilon + b^{\eta+2}(V+W)) \, . \tag{6.99}$$

Turning to $\tilde{F}$, by the integration bound (6.72),

$$|(\nabla^l \mathcal{K}^\lambda \tilde{F})_x| \lesssim |\!|\!| x |\!|\!|^{-|l|_{\mathfrak{s}}} b^{\eta+2}(V^* + W^*) \, . \tag{6.100}$$

Turning to Y and $\mathring{Y}$, since $\mathcal{L}(Y - \mathring{Y}) = 0$ on Ω_b, it follows from weighted Schauder estimates [Lie96, Thm. 4.9, p.59] and the maximum principle that, for any $|k|_{\mathfrak{s}} < 2$,

$$|\nabla^k (Y - \mathring{Y})_x| \lesssim |\!|\!| x |\!|\!|^{-|k|_{\mathfrak{s}}} \|Y - \mathring{Y}\|_{\infty;\partial\Omega_b} \, . \tag{6.101}$$

Furthermore, since $\varphi = \mathring{\varphi}$ on $\partial\Omega_b$ by our choice of $\mathring{\varphi}$,

$$\begin{aligned} \|Y - \mathring{Y}\|_{\infty;\partial\Omega_b} &= \big\|K^\lambda * \mathcal{R}\big\{P(\boldsymbol{\varphi})\mathbf{1} - P(\mathring{\boldsymbol{\varphi}})\mathbf{1} + \tilde{F}\big\}\big\|_{\infty;\partial\Omega_b} \\ &\lesssim b^{\eta+2}\big\{\|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2}(\varepsilon + b^{\eta+2}(V+W)) + (V^* + W^*)\big\} \end{aligned} \tag{6.102}$$

where the final bound is due (6.99)-(6.100). Treating $Y, \mathring{Y}$ as modelled distributions, we obtain from (6.101)-(6.102)

$$\|Y - \mathring{Y}\|_{\infty;<2} \lesssim b^{\eta+2}\big\{\|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2}(\varepsilon + b^{\eta+2}(V+W)) + (V^* + W^*)\big\} \, . \tag{6.103}$$

Finally, combining (6.99), (6.100), and (6.103), and using the definitions (6.98), we obtain

$$\|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} \lesssim b^{\eta+2}\big\{\|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2}(\varepsilon + b^{\eta+2}(V+W)) + (V^*+W^*)\big\} .$$

By (6.94) in Lemma 6.69, we have $V + W \le 1/2$. Taking b sufficiently small, we conclude that

$$\|\boldsymbol{\varphi} - \mathring{\boldsymbol{\varphi}}\|_{\infty;<2} \lesssim b^{\eta+2}(V^*+W^*) ,$$

which, by (6.95) in Lemma 6.69, can be made arbitrarily small by taking $r \to 0$. In particular, if $r > 0$ is sufficiently small, then $|\varphi(0) - \mathring{\varphi}(0)| < \delta$ and thus $|\varphi(0)| < \varepsilon - \delta$. ∎

Remark 6.71. The reason we require $\eta(\tau,\mathbf{k}), \bar\eta(\tau,\mathbf{k})$ to be non-negative in Assumption 6.22 is that, if these exponents were negative, we could not obtain (6.83). It would be natural in this case to follow the argument in Section 4.2.3 and consider the largest b such that $|\varphi| \in [\frac{\varepsilon}{2}, \varepsilon]$ on Ω_b, so the estimates in Section 6.9 hold on Ω_b. However, because b depends on φ, and it is possible that $|\mathring{\varphi}| \approx \varepsilon$ on a large subset of $\partial\Omega_b$, the domain Ω_b may be too small to ensure $|\mathring{\varphi}(0)| < \varepsilon - \delta$ for $\delta > 0$ *uniform* in φ, so the proof of the final Lemma 6.70 would not go through. For ODEs (Sections 4.2.3 and 5) this is not an issue because, by maximality of b and the fact that the boundary of the domain is a single point, we have $|\mathring{\varphi}| = \varepsilon/2$ on the boundary, thus $\mathring{\varphi}$ does not need 'space' to come down.

A Extension of distributions

Fix a scaling $\mathfrak{s} = (\mathfrak{s}_1, \dots, \mathfrak{s}_d) \in \mathbb{N}^d$ and follow the notation of Section 1.1.1. We use the shorthand $\mathfrak{B}_b = \mathfrak{B}(0,b)$. Fix $b \in (0,1]$ and, for $x \in \mathfrak{B}_b$, denote

$$[\![x]\!] = \inf_{y \in \mathrm{bd}(\mathfrak{B}_b)} |x-y|_{\mathfrak{s}} ,$$

where $\mathrm{bd}(A) \subset \mathbb{R}^d$ is the topological boundary of a set $A \subset \mathbb{R}^d$. For $\eta \le 0$, let $\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)$ denote the Banach space of distributions $\xi \in \mathcal{D}'(\mathfrak{B}_b)$ for which

$$[\![\xi]\!]_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} = \sup_{\psi\in\mathcal{B}^r} \sup_{x\in\mathfrak{B}_b} \sup_{0<\delta<\frac12[\![x]\!]} \delta^{-\eta}|\langle \xi, \psi_x^\delta\rangle| < \infty ,$$

where $r = \lfloor 1-\eta \rfloor \ge 1$ and we recall $\mathcal{B}^r$ from (1.6).

We start with a simple lemma, the proof of which is clear.

Lemma A.1. *Consider $C > 0$. An equivalent norm on $\mathcal{C}^\eta(\mathbb{R}^d)$ is*

$$\sup_{z\in\mathbb{R}^d} \sup_{\delta\in(0,1)} \sup_{\psi} \delta^{-\eta}|\langle \xi, \psi\rangle|$$

where the final sup *is taken over all*

$$\psi \in \mathcal{C}_c^\infty(\mathfrak{B}(z,\delta)) \quad \text{such that, for all } |k|_{\mathfrak{s}} \le r, \quad \|D^k\psi\|_\infty \le C\delta^{-|\mathfrak{s}|-|k|_{\mathfrak{s}}} . \tag{A.1}$$

The same statement holds for $\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)$ but with the constraints that $z \in \mathfrak{B}_b$ and $\mathfrak{B}(z,2\delta) \subset \mathfrak{B}_b$.

The following extension lemma is the main result of this appendix. Similar results can be found in the literature (see, e.g. [Ryc98, Ryc99] for extensions whose support goes outside $\mathrm{cl}\,\mathfrak{B}_b$), but we prefer to give a self-contained proof which, we believe, is comparatively simple.

Lemma A.2 (Extension)**.** *Let $\eta < 0$ with $\eta \notin \mathbb{Z}$. There exists a bounded linear operator $\mathcal{E}\colon \widehat{\mathcal{C}}^\eta(\mathfrak{B}_b) \to \mathcal{C}^\eta(\mathbb{R}^d)$ such that $\operatorname{supp}\mathcal{E}(\xi) \subset \operatorname{cl}\mathfrak{B}_b$ and $\langle \mathcal{E}(\xi), \psi\rangle = \langle \xi, \psi\rangle$ for all $\psi \in \mathcal{C}_c^\infty(\mathfrak{B}_b)$, and such that the operator norm of $\mathcal{E}$ does not depend on b.*

Proof. The key to the proof is a partition of unity $\{\varphi_x\}_{x\in\Lambda}$ such that

(a) Λ is a countable subset of $\mathfrak{B}_b$ and $\sum_{x\in\Lambda}\varphi_x = 1$ on $\mathfrak{B}_b$,

(b) $\varphi_x\colon \mathfrak{B}_b \to [0,1]$ with $\operatorname{supp}\varphi_x$ contained in the open ball $\mathfrak{B}(x, \frac{1}{2}[\![x]\!])$,

(c) $\|D^k\varphi_x\|_\infty \lesssim [\![x]\!]^{-|k|_\mathfrak{s}}$ uniformly in $|k|_\mathfrak{s} \le r$, $x \in \Lambda$, and $b \in (0,1]$, and

(d) for every $x \in \Lambda$, the number of $y \in \Lambda$ such that $\mathfrak{B}(x, \frac{1}{2}[\![x]\!]) \cap \mathfrak{B}(y, \frac{1}{2}[\![y]\!]) \neq \emptyset$ is bounded from above by a constant $N \ge 1$ that is uniform in x, y.

The existence of such a partition of unity follows in the same manner as in [Whi34] or [Ste70, Sec. VI.1], see also [Mar18, Sec. 5.3.3] and [CCHS22, Proof of Lem A.2].

Note that $\operatorname{bd}\mathfrak{B}_b$ is a box in which each side is parallel to a hyperplane $\{z : z_i = 0\}$ for some $i \in [d]$. For every $x \in \Lambda$, there exists a point $p_x \in \operatorname{bd}\mathfrak{B}_b$ such that $|x - p_x|_\mathfrak{s} \le 2[\![x]\!]$ and such that p_x is at least distance $\frac{1}{4}[\![x]\!]$ from the intersection of any two sides of $\mathfrak{B}_b$. Let ∂B_x be the side of $\operatorname{cl}\mathfrak{B}_b$ containing p_x and denote $\mathfrak{s}_x = \mathfrak{s}_i$ where $i \in [d]$ is such that ∂B_x is parallel to the hyperplane $\{z : z_i = 0\}$. Denote $B_x = \{y \in \operatorname{bd}\mathfrak{B}_b : |y - p_x|_\mathfrak{s} < \frac{1}{8}[\![x]\!]\}$. Remark that $B_x \subset \partial B_x$ by our choice of p_x. We let $⨍_{B_x} = \frac{1}{|B_x|}\int_{B_x}$ where the integral $\int_{B_x}$ and volume $|B_x|$ are taken with respect to the $(d-1)$-dimensional Lebesgue measure on B_x (so $|B_x| \asymp [\![x]\!]^{|\mathfrak{s}|-\mathfrak{s}_x}$).

Fix $\delta \in (0,1)$, $z \in \mathbb{R}^d$, and $\psi \in \mathcal{C}_c^\infty(\mathfrak{B}(z,\delta))$ as in (A.1) with $C = 1$. For every $x \in \Lambda$ we define $\psi_x \in \mathcal{C}_c^\infty(\mathfrak{B}(x, \frac{1}{2}[\![x]\!]))$ by

$$\psi_x(y) = \varphi_x(y)\Big\{\psi(y) - \sum_{0\le|k|_\mathfrak{s}<r-\mathfrak{s}_x} ⨍_{B_x} \frac{D^k\psi(p)}{k!}(y-p)^k\,\mathrm{d}p\Big\} \stackrel{\text{def}}{=} \varphi_x(y)\mathrm{T}_x^{r-\mathfrak{s}_x}(y)\,.$$

While $\operatorname{supp}\psi \subset \mathfrak{B}(z,\delta)$, remark that $\operatorname{supp}\mathrm{T}_x^{r-\mathfrak{s}_x}$ might not be contained in $\mathfrak{B}(z,\delta)$.

Fix furthermore $\xi \in \widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)$. We claim that the sum

$$\langle \mathcal{E}(\xi), \psi\rangle \stackrel{\text{def}}{=} \sum_{x\in\Lambda}\langle \xi, \psi_x\rangle \tag{A.2}$$

converges absolutely, agrees with $\langle \xi, \psi\rangle$ for $\psi \in \mathcal{C}_c^\infty(\mathfrak{B}_b)$, and satisfies the bound

$$|\langle \mathcal{E}(\xi), \psi\rangle| \lesssim \delta^\eta \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)}\,. \tag{A.3}$$

Clearly $\langle \mathcal{E}(\xi), \psi\rangle = 0$ whenever $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^d \setminus \operatorname{cl}\mathfrak{B}_b)$ and thus $\operatorname{supp}\mathcal{E}(\xi) \subset \operatorname{cl}\mathfrak{B}_b$. Therefore the proof follows from Lemma A.1 once the claim is shown.

To this end, assume first $\psi \in \mathcal{C}_c^\infty(\mathfrak{B}_b)$. Then the sum in (A.2) is finite by property (d). Moreover, $D^k\psi(p) = 0$ for all $p \in \operatorname{bd}\mathfrak{B}_b$ and therefore, by (a), $\sum_{x\in\Lambda}\psi_x = \psi$ and $\langle \mathcal{E}(\xi), \psi\rangle = \langle \xi, \psi\rangle$. Furthermore, if $\mathfrak{B}(z, 2\delta) \subset \mathfrak{B}_b$, then the claimed bound (A.3) follows from the equivalent norm on $\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)$ of Lemma A.1.

Suppose now that $\mathfrak{B}(z,2\delta)$ is not contained in $\mathfrak{B}_b$ (and we do not necessarily assume that ψ is supported in $\mathfrak{B}_b$). Consider first $[\![x]\!] \le 8\delta$. Let $l \in \mathbb{N}^d$ with $|l|_\mathfrak{s} \le r$. By (A.1), if $|l|_\mathfrak{s} < r - \mathfrak{s}_x$ and $r - \mathfrak{s}_x > 0$, then

$$\|D^l\mathrm{T}_x^{r-\mathfrak{s}_x}\|_{\infty;\mathfrak{B}(x,\frac{1}{2}[\![x]\!])} \lesssim \delta^{-|\mathfrak{s}|-r+\mathfrak{s}_x}[\![x]\!]^{r-\mathfrak{s}_x-|l|_\mathfrak{s}}\,,$$

while if $|l|_{\mathfrak{s}} \geq r - \mathfrak{s}_x$ (which is always the case if $r - \mathfrak{s}_x \leq 0$), then $D^l \mathrm{T}_x^{r-\mathfrak{s}_x} = D^l \psi$ and thus

$$\|D^l \mathrm{T}_x^{r-\mathfrak{s}_x}\|_\infty \lesssim \delta^{-|\mathfrak{s}|-|l|_{\mathfrak{s}}} .$$

Then, for any multi-index $k \in \mathbb{N}^d$ with $|k|_{\mathfrak{s}} \leq r$, by (b)-(c),

$$\begin{aligned}\|D^k \psi_x\|_\infty &\lesssim \sum_{l \leq k,\, |l|_{\mathfrak{s}} < r - \mathfrak{s}_x} [\![x]\!]^{-|k-l|_{\mathfrak{s}}} \delta^{-|\mathfrak{s}|-r+\mathfrak{s}_x} [\![x]\!]^{r-\mathfrak{s}_x-|l|_{\mathfrak{s}}} + \sum_{l \leq k,\, |l|_{\mathfrak{s}} \geq r - \mathfrak{s}_x} [\![x]\!]^{-|k-l|_{\mathfrak{s}}} \delta^{-|\mathfrak{s}|-|l|_{\mathfrak{s}}} \\ &\lesssim \begin{cases} [\![x]\!]^{r-\mathfrak{s}_x-|k|_{\mathfrak{s}}} \delta^{-|\mathfrak{s}|-r+\mathfrak{s}_x} & \text{if } r - \mathfrak{s}_x > 0 \\ [\![x]\!]^{-|k|_{\mathfrak{s}}} \delta^{-|\mathfrak{s}|} & \text{if } r - \mathfrak{s}_x \leq 0 . \end{cases}\end{aligned}$$

It follows that

$$\max_{|k|_{\mathfrak{s}} \leq r} [\![x]\!]^{|\mathfrak{s}|+|k|_{\mathfrak{s}}} \|D^k \psi_x\|_\infty \lesssim \begin{cases} [\![x]\!]^{r-\mathfrak{s}_x+|\mathfrak{s}|} \delta^{-|\mathfrak{s}|-r+\mathfrak{s}_x} & \text{if } r - \mathfrak{s}_x > 0 \\ [\![x]\!]^{|\mathfrak{s}|} \delta^{-|\mathfrak{s}|} & \text{if } r - \mathfrak{s}_x \leq 0 \end{cases}$$

and therefore, since ψ_x is supported in $\mathfrak{B}(x, \frac{1}{2}[\![x]\!])$, it follows from Lemma A.1 that

$$|\langle \xi, \psi_x \rangle| \lesssim [\![x]\!]^{\eta+|\mathfrak{s}|+(r-\mathfrak{s}_x)\vee 0} \delta^{-|\mathfrak{s}|-(r-\mathfrak{s}_x)\vee 0} \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} .$$

By (d), uniformly in $n \geq 0$, there are $O(2^{n(|\mathfrak{s}|-\mathfrak{s}_i)} \delta^{|\mathfrak{s}|-\mathfrak{s}_i})$ points $x \in \Lambda$ in any ball of radius δ for the scaling $\mathfrak{s}$ such that $\mathfrak{s}_x = \mathfrak{s}_i$ and $[\![x]\!] \in [2^{-n}, 2^{-n+1}]$. Moreover, if $[\![x]\!] \leq 8\delta$ and x is outside the ball $\mathfrak{B}(z, C\delta)$ for $C > 0$ sufficiently large, then ψ vanishes on a neighbourhood of B_x and therefore $\psi_x = 0$. It follows that

$$\sum_{[\![x]\!] \leq 8\delta} |\langle \xi, \psi_x \rangle| \lesssim \max_{i \in [d]} \sum_{2^{-n} \leq C\delta} 2^{n(|\mathfrak{s}|-\mathfrak{s}_i-\eta-|\mathfrak{s}|-(r-\mathfrak{s}_i)\vee 0)} \delta^{|\mathfrak{s}|-\mathfrak{s}_i-|\mathfrak{s}|-(r-\mathfrak{s}_i)\vee 0} \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} \lesssim \delta^\eta \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} , \tag{A.4}$$

where we use that the exponent of 2^n is strictly negative.

Consider now $[\![x]\!] > 8\delta$. Note that ψ vanishes on $\mathfrak{B}(x, \frac{1}{2}[\![x]\!])$ and thus, for $|l|_{\mathfrak{s}} \geq r - \mathfrak{s}_x$, we have $D^l \mathrm{T}_x^{r-\mathfrak{s}_x} \equiv 0$ on $\mathfrak{B}(x, \frac{1}{2}[\![x]\!])$. On the other hand, for $|l|_{\mathfrak{s}} < r - \mathfrak{s}_x$, by (A.1) and the fact that $|B_x|^{-1} \lesssim [\![x]\!]^{-|\mathfrak{s}|+\mathfrak{s}_x}$ and the support of ψ in B_x has volume at most $\lesssim \delta^{|\mathfrak{s}|-\mathfrak{s}_x}$,

$$\|D^l \mathrm{T}_x^{r-\mathfrak{s}_x}\|_{\infty;\mathfrak{B}(x,\frac{1}{2}[\![x]\!])} \lesssim [\![x]\!]^{-|\mathfrak{s}|+\mathfrak{s}_x} \delta^{|\mathfrak{s}|-\mathfrak{s}_x} \delta^{-|\mathfrak{s}|-r+\mathfrak{s}_x+1} [\![x]\!]^{r-\mathfrak{s}_x-|l|_{\mathfrak{s}}-1} = \delta^{-r+1} [\![x]\!]^{-|\mathfrak{s}|+r-1-|l|_{\mathfrak{s}}} .$$

Therefore, if $r - \mathfrak{s}_x \leq 0$, then $\psi_x \equiv 0$, and we are done by the previous bound (A.4). Suppose now that $r - \mathfrak{s}_x > 0$. Then, using in addition (b)-(c), uniformly in $|k|_{\mathfrak{s}} \leq r$,

$$\|D^k \psi_x\|_\infty \lesssim \sum_{l \leq k,\, |l|_{\mathfrak{s}} < r - \mathfrak{s}_x} [\![x]\!]^{-|k-l|_{\mathfrak{s}}} \delta^{-r+1} [\![x]\!]^{-|\mathfrak{s}|+r-1-|l|_{\mathfrak{s}}} \lesssim [\![x]\!]^{-|\mathfrak{s}|+r-1-|k|_{\mathfrak{s}}} \delta^{-r+1} .$$

It follows that

$$\max_{|k|_{\mathfrak{s}} \leq r} [\![x]\!]^{|\mathfrak{s}|+|k|_{\mathfrak{s}}} \|D^k \psi_x\|_\infty \lesssim [\![x]\!]^{r-1} \delta^{-r+1} .$$

Since ψ_x is supported in $\mathfrak{B}(x, \frac{1}{2}[\![x]\!])$, we obtain from Lemma A.1 that

$$|\langle \xi, \psi_x \rangle| \lesssim [\![x]\!]^{\eta+r-1} \delta^{-r+1} \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} .$$

By (d), uniformly in $n \geq 0$, there are $O(1)$ points $x \in \Lambda$ such that $[\![x]\!] > 8\delta$, $[\![x]\!] \in [2^{-n}, 2^{-n+1}]$, and B_x overlaps with $\mathfrak{B}(z, \delta)$. It follows that

$$\sum_{[\![x]\!] \geq 8\delta} |\langle \xi, \psi_x \rangle| \lesssim \max_{i \in [d]} \sum_{2^{-n} \geq \delta} 2^{n(-\eta-r+1)} \delta^{-r+1} \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} \lesssim \delta^\eta \|\xi\|_{\widehat{\mathcal{C}}^\eta(\mathfrak{B}_b)} ,$$

where we used $-\eta - r + 1 > 0$ because we assumed $\eta \notin \mathbb{Z}$. This proves (A.3) and the claim. ∎

Remark A.3. Note that $\operatorname{supp} \xi \subset \operatorname{supp} \mathcal{E}(\xi)$ and that the containment is strict in general. It is not clear to us whether an extension exists that does not enlarge the support.

Acknowledgments

I. C. acknowledges support from the European Research Council via the grant SQGT 101116964. M. G. was partly supported by UK Research and Innovation via the grant "StochFields" EP/Z534328/1.

References

[BC19] H. Boedihardjo and I. Chevyrev. An isomorphism between branched and geometric rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**, no. 2, (2019), 1131–1148. doi:10.1214/18-aihp912.

[BC24a] B. Bringmann and S. Cao. Global well-posedness of the dynamical sine-Gordon model up to 6π. *arXiv e-prints* (2024). To appear in *Annals of Probability*. arXiv:2410.15493.

[BC24b] B. Bringmann and S. Cao. Global well-posedness of the stochastic Abelian-Higgs equations in two dimensions. *arXiv e-prints* (2024). arXiv:2403.16878.

[BCCH21] Y. Bruned, A. Chandra, I. Chevyrev, and M. Hairer. Renormalising SPDEs in regularity structures. *J. Eur. Math. Soc. (JEMS)* **23**, no. 3, (2021), 869–947. doi:10.4171/jems/1025.

[BCFP19] Y. Bruned, I. Chevyrev, P. K. Friz, and R. Preiss. A rough path perspective on renormalization. *J. Funct. Anal.* **277**, no. 11, (2019), 108283, 60. doi:10.1016/j.jfa.2019.108283.

[BCMW22] T. Bonnefoi, A. Chandra, A. Moinat, and H. Weber. A priori bounds for rough differential equations with a non-linear damping term. *J. Differential Equations* **318**, (2022), 58–93. doi:10.1016/j.jde.2022.02.006.

[BCZ24] L. Broux, F. Caravenna, and L. Zambotti. Hairer's multilevel Schauder estimates without regularity structures. *Trans. Amer. Math. Soc.* **377**, no. 10, (2024), 6981–7035. doi:10.1090/tran/9245.

[BGHZ22] Y. Bruned, F. Gabriel, M. Hairer, and L. Zambotti. Geometric stochastic heat equations. *J. Amer. Math. Soc.* **35**, no. 1, (2022), 1–80. doi:10.1090/jams/977.

[BH20] I. Bailleul and M. Hoshino. A tourist's guide to regularity structures and singular stochastic PDEs. *arXiv e-prints* (2020). To appear in EMS Surv. Math. Sci. arXiv:2006.03524. doi:10.4171/EMSS/87.

[BHZ19] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. *Invent. Math.* **215**, no. 3, (2019), 1039–1156. doi:10.1007/s00222-018-0841-x.

[BK23] Y. Bruned and F. Katsetsiadis. Post-Lie algebras in regularity structures. *Forum Math. Sigma* **11**, (2023), Paper No. e98, 20. doi:10.1017/fms.2023.93.

[BM23] Y. Bruned and D. Manchon. Algebraic deformation for (S)PDEs. *J. Math. Soc. Japan* **75**, no. 2, (2023), 485–526. arXiv:2011.05907. doi:10.2969/jmsj/88028802.

[BOS25] L. Broux, F. Otto, and R. Steele. Multi-index Based Solution Theory to the Φ^4 Equation in the Full Subcritical Regime. *arXiv e-prints* (2025). arXiv:2503.01621.

[CCHS22] A. Chandra, I. Chevyrev, M. Hairer, and H. Shen. Langevin dynamic for the 2D Yang-Mills measure. *Publ. math. IHÉS* (2022). arXiv:2006.04987. doi:10.1007/s10240-022-00132-0.

[CCHS24] A. Chandra, I. Chevyrev, M. Hairer, and H. Shen. Stochastic quantisation of Yang-Mills-Higgs in 3D. *Invent. Math.* **237**, no. 2, (2024), 541–696. doi:10.1007/s00222-024-01264-2.

[CdW24] A. Chandra, G. de Lima Feltes, and H. Weber. A priori bounds for 2-d generalised Parabolic Anderson Model. *arXiv e-prints* (2024). arXiv:2402.05544.

[CH16] A. Chandra and M. Hairer. An analytic BPHZ theorem for regularity structures. *ArXiv e-prints* (2016). arXiv:1612.08138.

[Che22] I. Chevyrev. Hopf and pre-Lie algebras in regularity structures. *arXiv e-prints* (2022). arXiv:2206.14557.

[Che25] I. Chevyrev. Rough path theory. In R. Szabo and M. Bojowald, eds., *Encyclopedia of Mathematical Physics (Second Edition)*, 1–15. Academic Press, Oxford, second edition ed., 2025. doi:https://doi.org/10.1016/B978-0-323-95703-8.00027-6.

[CMW23] A. Chandra, A. Moinat, and H. Weber. A priori bounds for the Φ^4 equation in the full sub-critical regime. *Arch. Ration. Mech. Anal.* **247**, no. 3, (2023), Paper No. 48, 76. doi:10.1007/s00205-023-01876-7.

[CS23] I. Chevyrev and H. Shen. Invariant measure and universality of the 2D Yang-Mills Langevin dynamic. *arXiv e-prints* (2023). arXiv:2302.12160.

[Dav08] A. M. Davie. Differential equations driven by rough paths: An approach via discrete approximation. *Applied Mathematics Research eXpress* **2008**, (2008), abm009. doi:10.1093/amrx/abm009.

[DGR23] P. Duch, M. Gubinelli, and P. Rinaldi. Parabolic stochastic quantisation of the fractional Φ^4_3 model in the full subcritical regime. *arXiv e-prints* (2023). arXiv:2303.18112.

[DPD03] G. Da Prato and A. Debussche. Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31**, no. 4, (2003), 1900–1916. doi:10.1214/aop/1068646370.

[Duc25] P. Duch. Flow equation approach to singular stochastic PDEs. *Probab. Math. Phys.* **6**, no. 2, (2025), 327–437. doi:10.2140/pmp.2025.6.327.

[EW24] S. Esquivel and H. Weber. A priori bounds for the dynamic fractional Φ^4 model on $\mathbb{T}^3$ in the full subcritical regime. *arXiv e-prints* (2024). arXiv:2411.16536.

[FdLP06] D. Feyel and A. de La Pradelle. Curvilinear integrals along enriched paths. *Electron. J. Probab.* **11**, (2006), no. 34, 860–892. doi:10.1214/EJP.v11-356.

[FH20] P. K. Friz and M. Hairer. *A course on rough paths.* Universitext. Springer, Cham, [2020] ©2020, xvi+346. With an introduction to regularity structures, Second edition of [3289027]. doi:10.1007/978-3-030-41556-3.

[FV08] P. Friz and N. Victoir. Euler estimates for rough differential equations. *J. Differential Equations* **244**, no. 2, (2008), 388–412. doi:10.1016/j.jde.2007.10.008.

[FV10] P. K. Friz and N. B. Victoir. *Multidimensional stochastic processes as rough paths*, vol. 120 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2010, xiv+656. Theory and applications. doi:10.1017/CBO9780511845079.

[GH19] M. Gerencsér and M. Hairer. Singular SPDEs in domains with boundaries. *Probab. Theory Related Fields* **173**, no. 3-4, (2019), 697–758. doi:10.1007/s00440-018-0841-1.

[GH21] M. Gubinelli and M. Hofmanová. A PDE construction of the Euclidean φ^4_3 quantum field theory. *Comm. Math. Phys.* **384**, no. 1, (2021), 1–75. doi:10.1007/s00220-021-04022-0.

[GIP15] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, (2015), e6, 75. arXiv:1210.2684v3. doi:10.1017/fmp.2015.2.

[Gub04] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.* **216**, no. 1, (2004), 86–140. doi:10.1016/j.jfa.2004.01.002.

[Gub10] M. Gubinelli. Ramification of rough paths. *J. Differential Equations* **248**, no. 4, (2010), 693–721. doi:10.1016/j.jde.2009.11.015.

[Hai14] M. Hairer. A theory of regularity structures. *Invent. Math.* **198**, no. 2, (2014), 269–504. arXiv:1303.5113. doi:10.1007/s00222-014-0505-4.

[Hai15] M. Hairer. Introduction to regularity structures. *Braz. J. Probab. Stat.* **29**, no. 2, (2015), 175–210. doi:10.1214/14-BJPS241.

[HK15] M. Hairer and D. Kelly. Geometric versus non-geometric rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.* **51**, no. 1, (2015), 207–251. doi:10.1214/13-AIHP564.

[HR24] M. Hairer and T. Rosati. Global existence for perturbations of the 2D stochastic Navier-Stokes equations with space-time white noise. *Ann. PDE* **10**, no. 1, (2024), Paper No. 3, 46. doi:10.1007/s40818-023-00165-6.

[HS22] M. Hairer and R. Steele. The Φ^4_3 measure has sub-Gaussian tails. *J. Stat. Phys.* **186**, no. 3, (2022), Paper No. 38, 25. doi:10.1007/s10955-021-02866-3.

[HS24] M. Hairer and R. Steele. The BPHZ theorem for regularity structures via the spectral gap inequality. *Arch. Ration. Mech. Anal.* **248**, no. 1, (2024), Paper No. 9, 81. arXiv:2301.10081. doi:10.1007/s00205-023-01946-w.

[HZ25] M. Hairer and W. Zhao. Ergodicity of 2D singular stochastic Navier–Stokes equations. *Probab. Math. Phys.* **6**, no. 3, (2025), 777–818. doi:10.2140/pmp.2025.6.777.

[JP25] R. Jin and N. Perkowski. The compact support property of rough super Brownian motion on $\mathbb{R}^2$. *Stochastic Process. Appl.* **182**, (2025), Paper No. 104568, 24. doi:10.1016/j.spa.2025.104568.

[Kup16] A. Kupiainen. Renormalization group and stochastic PDEs. *Annales Henri Poincaré* **17**, no. 3, (2016), 497–535. arXiv:1410.3094. doi:10.1007/s00023-015-0408-y.

[Lie96] G. M. Lieberman. *Second order parabolic differential equations*. World Scientific Publishing Co., Inc., River Edge, NJ, 1996, xii+439. doi:10.1142/3302.

[LOT23] 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. doi:10.1090/cams/16.

[LOTT24] P. Linares, F. Otto, M. Tempelmayr, and P. Tsatsoulis. A diagram-free approach to the stochastic estimates in regularity structures. *Invent. Math.* **237**, no. 3, (2024), 1469–1565. doi:10.1007/s00222-024-01275-z.

[Lyo94] T. Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.* **1**, no. 4, (1994), 451–464. doi:10.4310/MRL.1994.v1.n4.a5.

[Lyo98] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14**, no. 2, (1998), 215–310. doi:10.4171/RMI/240.

[Mar18] J. Martin. *Refinements of the Solution Theory for Singular SPDEs*. Ph.D. thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2018. doi:10.18452/19329.

[MW17] J.-C. Mourrat and H. Weber. Global well-posedness of the dynamic Φ^4 model in the plane. *Ann. Probab.* **45**, no. 4, (2017), 2398–2476. doi:10.1214/16-AOP1116.

[MW20a] A. Moinat and H. Weber. Local bounds for stochastic reaction diffusion equations. *Electron. J. Probab.* **25**, (2020), Paper No. 17, 26. doi:10.1214/19-ejp397.

[MW20b] A. Moinat and H. Weber. Space-time localisation for the dynamic Φ^4_3 model. *Comm. Pure Appl. Math.* **73**, no. 12, (2020), 2519–2555. doi:10.1002/cpa.21925.

[OSSW25] F. Otto, J. Sauer, S. A. Smith, and H. Weber. A priori bounds for quasilinear SPDEs in the full subcritical regime. *J. Eur. Math. Soc. (JEMS)* **27**, no. 1, (2025), 71–118. doi:10.4171/jems/1574.

[OW19] F. Otto and H. Weber. Quasilinear SPDEs via rough paths. *Arch. Ration. Mech. Anal.* **232**, no. 2, (2019), 873–950. doi:10.1007/s00205-018-01335-8.

[Ryc98] V. S. Rychkov. Intrinsic characterizations of distribution spaces on domains. *Studia Math.* **127**, no. 3, (1998), 277–298. doi:10.4064/sm-127-3-277-298.

[Ryc99] V. S. Rychkov. On restrictions and extensions of the Besov and Triebel-Lizorkin spaces with respect to Lipschitz domains. *J. London Math. Soc. (2)* **60**, no. 1, (1999), 237–257. doi:10.1112/S0024610799007723.

[Sim97] L. Simon. Schauder estimates by scaling. *Calc. Var. Partial Differential Equations* **5**, no. 5, (1997), 391–407. doi:10.1007/s005260050072.

[SS25] J. Sauer and S. A. Smith. Schauder estimates for germs by scaling. *Stochastics and Partial Differential Equations: Analysis and Computations* (2025). arXiv:2412.01486. doi:10.1007/s40072-025-00361-4.

[Ste70] E. M. Stein. *Singular integrals and differentiability properties of functions*. Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, NJ, 1970, xiv+290. doi:10.1515/9781400883882.

[SZZ25] H. Shen, R. Zhu, and X. Zhu. Global well-posedness for 2d generalized parabolic anderson model via paracontrolled calculus. *Stochastics and Partial Differential Equations: Analysis and Computations* (2025). arXiv:2402.05544. doi:10.1007/s40072-025-00358-z.

[Whi34] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36**, no. 1, (1934), 63–89. doi:10.2307/1989708.

[ZK23] P. Zorin-Kranich. The reconstruction theorem in quasinormed spaces. *Rev. Mat. Iberoam.* **39**, no. 4, (2023), 1233–1246. doi:10.4171/rmi/1355.