

# A Littlewood-Paley description of modelled distributions

Jörg Martin\*

Institut für Mathematik  
Humboldt–Universität zu Berlin  
martin@math.hu-berlin.de

Nicolas Perkowski†

Max-Planck-Institute for Mathematics in the Sciences, Leipzig  
& Humboldt–Universität zu Berlin  
perkowski@math.hu-berlin.de

August 1, 2018

## Abstract

We exhibit a fundamental link between Hairer’s theory of regularity structures [Hai14] and the paracontrolled calculus of [GIP15]. By using paraproducts we provide a Littlewood-Paley description of the spaces of modelled distributions in regularity structures that is similar to the Besov description of classical Hölder spaces.

## 1 Introduction

This article builds a bridge between two different approaches that arose from the study of singular stochastic partial differential equations (singular SPDEs). These equations are ill posed because of the interplay of highly irregular noise with nonlinearities, which often leads to resonances that may have to be removed by a renormalization procedure. A new way of describing distributions, their regularity, and operations on them was required in order to give a meaningful solution theory for singular SPDEs, and this was implemented differently by different groups. Our aim is to reveal a link between *regularity structures*, which were developed by Hairer in [Hai14] and applied abundantly since (for example [HQ15, HL18, HS17, HP15, CFG17, HM18]), and the Fourier approach of *paracontrolled distributions*, which first appeared in the work [GIP15] by Gubinelli, Imkeller and Perkowski and was used for instance in [CC18b, CC18a, MW17a, GP17, AC15, ZZ18, MP17]. These two techniques are by no means the only tools for singular SPDEs and alternative views were provided by Kupiainen in his application of renormalization group techniques [Kup16, KM17], and by Otto and Weber with their rough path flavored approach [OW16, OSSW18].

In this paper however we will only focus on the comparison of two seemingly distinct notions from regularity structures and paracontrolled distributions. Central to the theory of [Hai14] is

---

\*Financial support by the DFG via Research Training Group RTG 1845 is gratefully acknowledged.

†Financial support by the DFG via the Heisenberg Program and via Research Unit FOR 2402 is gratefully acknowledged.

the concept of a *modelled distribution*. These are generalized Hölder functions that are classified by spaces called  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  with  $\gamma \in \mathbb{R}$ . The sets  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  collect functions  $F: \mathbb{R}^d \rightarrow \mathcal{T}$  that take values in a graded vector space  $\mathcal{T} = \bigoplus_{\alpha \in A} \mathcal{T}_\alpha$  (with  $A \subseteq \mathbb{R}$  and with  $\mathcal{T}_\alpha$  being normed spaces) and whose components  $F^\alpha: \mathbb{R}^d \rightarrow \mathcal{T}_\alpha$  satisfy for  $x, y \in \mathbb{R}^d$ <sup>1</sup>

$$\|F_y^\alpha - \Gamma_{yx}^\alpha F_x\|_{\mathcal{T}_\alpha} \lesssim |y - x|^{\gamma - \alpha} \quad (1)$$

for  $\alpha < \gamma$ . The object  $\Gamma_{yx}^\alpha F_x \in \mathcal{T}_\alpha$  in (1) should be thought of as a local ‘‘Taylor-like’’ expansion. The most basic case arises when we consider the so called polynomial regularity structure  $\overline{\mathcal{T}}$ , where  $F^\alpha$  just represents the derivative of order  $\alpha$  of a function  $f$  and  $\Gamma_{yx}^\alpha F_x$  incorporates the Taylor expansion of this derivative, so that  $\mathcal{D}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  just represents the  $\gamma$ -Hölder continuous functions (for a precise formulation see Lemma 2.23 below). It is rather classical that the space of  $\gamma$ -Hölder continuous functions can also be described via Littlewood-Paley theory as a Besov space  $\mathcal{B}^\gamma(\mathbb{R}^d)$ , so that there are in fact various ‘‘classical’’ descriptions of  $\mathcal{D}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$ . However, in general the space  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  can be of a far more complex structure and usually it involves expansions in non-polynomial objects, which are typically constructed from an underlying SPDE. In such cases there is no description of  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  in terms of more familiar function spaces. The central role of modelled distributions in the theory of regularity structure is due to the so called reconstruction theorem [Hai14, Theorem 3.10] which shows that there is a (unique) distribution  $f \in \mathcal{S}'(\mathbb{R}^d)$  that is described by the modelled distribution  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$ . On the level of  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  there is a robust theory for operations such as multiplication whose execution is typically hazardous on  $\mathcal{S}'(\mathbb{R}^d)$ , and describing the objects in an SPDE via modelled distributions leads to a robust formulation of the Schauder theory. But since there is no description of  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  in terms of classical function spaces, all these results have to be derived from scratch and are typically quite cumbersome to prove.

The linchpin in the paracontrolled framework is the notion of a *paracontrolled distribution*. Roughly speaking, a distribution  $f \in \mathcal{S}'(\mathbb{R}^d)$ , often the solution to an SPDE, is called paracontrolled if it can be smoothed by the subtraction of a *paraproduct*, a sort of ‘‘frequency modulation’’ of a given reference distribution. This allows us to transform the considered SPDE into an equation that can be solved via classical Schauder theory. Although the paracontrolled approach might seem quite different from the one in regularity structures, note that an increase of smoothness could be interpreted as the cancellation of fluctuations on small scales. From this point of view the paraproduct seems to capture the local behaviour of the considered distribution. This philosophy seems reminiscent to the idea of a local expansion in (1). Moreover, often similar objects appear in the expansion  $\Gamma_{yx}^\alpha F_x$  in (1) and the paraproducts used to control the solutions to SPDEs: Compare for instance the solution theory for the parabolic Anderson model [Hai14, GIP15], the  $\Phi_3^4$  model [Hai14, CC18b] or the KPZ equation [FH14, GP17]. Based on this similarity it has been conjectured in [GIP15] that there might be a one-to-one correspondence between modelled distributions and paracontrolled distributions. Here we give a precise formulation of that conjecture and we prove it.

More precisely, we propose the following description of a modelled distribution via paraproducts: In Definition 3.1 we define the paraproducts  $P(F, \Gamma^\alpha)$  for  $\alpha \in A$  and  $F: \mathbb{R}^d \rightarrow \mathcal{T} = \bigoplus_{\alpha \in A} \mathcal{T}_\alpha$ . We then say that  $F$  is in the space  $\mathcal{B}^\gamma(\mathbb{R}^d; \mathcal{T})$  if for  $\alpha < \gamma$

$$F^\alpha - P(F, \Gamma^\alpha) \in \mathcal{B}^{\gamma - \alpha}, \quad (2)$$

---

<sup>1</sup>Instead of  $|y - x|$  we will consider a scaled distance  $\|y - x\|_s$  below, as in [Hai14].

where  $\mathcal{B}^{\gamma-\alpha} = B_{\infty,\infty}^{\gamma-\alpha}$  denotes a Besov space, and if further a certain *structure condition* (Definition 3.4 below) is satisfied. In Theorem 4.1 we then show that

**Theorem.** For  $\gamma \in \mathbb{R} \setminus (\mathbb{N} + A)$  we have

$$\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) = \mathcal{B}^\gamma(\mathbb{R}^d; \mathcal{T}) \quad (3)$$

with equivalent norms.

One might consider  $\mathcal{B}^\gamma(\mathbb{R}^d; \mathcal{T})$  as a type of Besov space on  $\mathcal{T}$ . Indeed, in the case of the polynomial regularity structure  $\mathcal{T} = \overline{\mathcal{T}}$  it turns out that  $\mathcal{B}^\gamma(\mathbb{R}^d; \mathcal{T})$  just collapses to the classical Besov space  $\mathcal{B}^\gamma$ , so that (3) can be interpreted as a generalization of the Besov description of Hölder continuous functions. Since both,  $\mathcal{B}^\gamma$  and  $P(F, \Gamma^\alpha)$ , will be built using Littlewood-Paley theory, (3) is a Littlewood-Paley description of  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  and moreover a characterization in terms of more classical spaces from analysis.

Although we consider the theorem above an interesting observation in its own right, one should also judge its importance from a more practical point of view. For many operators, such as Fourier multipliers, it is more natural to work with the spectral description of function spaces provided by the Littlewood-Paley theory. The theorem above therefore yields a first insight on how to describe modelled distributions in this formulation. However, an elaboration of a full, say, Schauder-like statement based on (3) would still require a lot of technical work, which is why we refrain from doing so in this paper. For ideas and results in that direction we refer to the Ph.D. thesis of the first author [Mar18].

## Structure of this article

Section 2 recalls the fundamental concepts we will use. These are mostly some definitions and basic facts concerning Besov spaces and some elementary definitions from the theory of regularity structures. Since statements about Besov regularity, such as (2), actually concern objects with values in Banach spaces, we will also repeat the essentials of vector-valued distributions.

Conceptually new definitions are given in Section 3, where we introduce the paraproducts  $P(F, \Gamma^\alpha)$  and the space  $\mathcal{B}^\gamma(\mathbb{R}^d; \mathcal{T})$  mentioned above.

The main result of this paper, that is identity (3), is presented and proved in Section 4.

## Notation

We will write  $x \lesssim y$  to denote an inequality of the form  $x \leq C \cdot y$  for some constant  $C > 0$ , which is independent of  $x$  and  $y$ . We also use the notation  $x \gtrsim y$  to describe the relation  $y \lesssim x$ . We write  $x \approx y$  to indicate that both statements,  $x \lesssim y$  and  $x \gtrsim y$ , are true.

For indices  $i, j \in \mathbb{Z}$  we will write

$$i \lesssim j \quad \Leftrightarrow \quad i \leq j + N \quad (4)$$

where  $N \in \mathbb{Z}$  is a deterministic constant which is independent of  $i$  and  $j$ , i.e. if  $2^i \lesssim 2^j$ . We also use  $i \gtrsim j$  to denote  $i \lesssim j$  and  $i \sim j$  whenever  $i \lesssim j$  and  $j \lesssim i$ . The symbol  $\mathbb{N} = \{0, 1, 2, \dots\}$  denotes in this article the natural numbers including 0. The set  $\mathbb{N}_\times := \mathbb{N} \setminus \{0\} = \{1, 2, 3, \dots\}$  is then defined by exclusion of the latter.

We use the symbol  $C^r(\mathbb{R}^d)$  for the set of  $r \in \mathbb{N}$  times differentiable, complex-valued functions and set  $C^\infty(\mathbb{R}^d) = \bigcap_{r \in \mathbb{N}} C^r(\mathbb{R}^d)$ . To indicate bounded derivatives we add an index  $b$ , so that

$C_b(\mathbb{R}^d)$  denotes for instance the set of bounded, continuous and complex-valued functions, equipped with the norm

$$\|f\|_{C_b(\mathbb{R}^d)} = \sup_{x \in \mathbb{R}^d} |f(x)|.$$

$C_c^\infty(\mathbb{R}^d)$  is used for the subset of functions in  $C^\infty(\mathbb{R}^d)$  that have compact support.  $\mathcal{S}(\mathbb{R}^d)$  denotes the Schwartz space. A codomain of a function space which is different from the complex numbers will be mentioned in the notation via a semicolon, so that  $C^\infty(\mathbb{R}^d; [0, 1])$  denotes the smooth functions on  $\mathbb{R}^d$  which take values in the unit interval and  $C_c^\infty(\mathbb{R}^d; X)$  is the set of smooth, compactly supported functions with values in a Banach space  $X$ . Banach-valued distributions  $\mathcal{S}'(\mathbb{R}^d; X)$  and Besov spaces  $\mathcal{B}_s^r(\mathbb{R}^d; X)$  will be introduced in Subsection 2.1. Given a function  $f$  on  $\mathbb{R}^d$  we will occasionally write shorthand  $f_x := f(x)$  for  $x \in \mathbb{R}^d$  and use the “physics convention” in writing the differential right after the integral sign, for instance

$$\int_{\mathbb{R}^d} dx f_x.$$

We follow [Hai14] in the definition of anisotropic distances and scalings: Given a *scaling vector*

$$\mathfrak{s} \in \mathbb{N}_>^d, \quad (5)$$

we define for  $x \in \mathbb{R}^d$  the *scaled “norm”*

$$\|x\|_{\mathfrak{s}} := \sum_{i=1}^d |x_i|^{1/\mathfrak{s}_i}, \quad (6)$$

and we define the scaled unit ball as

$$B_{\mathfrak{s}}(0, 1) = \{x \in \mathbb{R}^d \mid \|x\|_{\mathfrak{s}} < 1\}.$$

For multi-indices  $k \in \mathbb{N}^d$  we set

$$|k|_{\mathfrak{s}} = \sum_{i=1}^d k_i \mathfrak{s}_i, \quad |\mathfrak{s}| = \sum_{i=1}^d \mathfrak{s}_i.$$

Finally we write for positive  $a > 0$

$$a^{\mathfrak{s}} := \text{diag}(a^{\mathfrak{s}_1}, a^{\mathfrak{s}_2}, \dots, a^{\mathfrak{s}_d}), \quad (7)$$

where  $\text{diag}(\cdot)$  denotes the diagonal matrix with diagonal “ $\cdot$ ”. For  $\varphi \in L^1(\mathbb{R}^d)$  we sometimes need the notation  $\varphi^\lambda$  by which we mean the  $L^1$ -scaling

$$\varphi^\lambda := \lambda^{-|\mathfrak{s}|} \varphi(\lambda^{-\mathfrak{s}} \cdot),$$

so that  $\varphi_x^\lambda$  should be read as

$$\varphi_x^\lambda = \lambda^{-|\mathfrak{s}|} \varphi(\lambda^{-\mathfrak{s}} x), \quad (8)$$

where the *matrix*  $\lambda^{-\mathfrak{s}}$  is defined as in 7.

**Remark 1.1.**  $\triangleleft$  Note that we slightly differ here from [Hai14], where this notation denotes the function  $z \mapsto \lambda^{-|\mathfrak{s}|} \varphi(\lambda^{-\mathfrak{s}}(z - x))$  instead. For us  $\varphi_x^\lambda$  denotes a complex number and not a function.

## 2 Background

In this section we recall the basic definitions we will make use of in this article. In Subsection 2.1 we repeat the definition and basic properties of Besov spaces. We need two slightly unusual modifications which is the allowance for *anisotropic* smoothness as well as the possibility for the considered distributions to take values in a Banach space. We therefore start with a short repetition concerning vector-valued distributions.

In Subsection 2.2 we recall a few elementary definitions from the theory of regularity structures. Essential to us is Definition 2.20 which introduces the space  $\mathcal{D}'(\mathbb{R}^d; \mathcal{T})$  of modelled distributions.

### 2.1 Anisotropic Besov spaces

#### Banach-valued distributions

It will be convenient for us to work with vector-valued distributions in this article. Although this theory is rather classical in the literature (see e.g. [ST87, Tre75] or even [Sch50]), the scalar-valued case is usually a more familiar concept, so that we here briefly repeat the most important aspects of the vector modification of distribution theory.

For a Banach space  $X$  we define the *Banach-valued* Schwartz distributions  $\mathcal{S}'(\mathbb{R}^d; X)$  as the set of continuous linear functionals

$$f : \mathcal{S}(\mathbb{R}^d) \rightarrow X,$$

where  $\mathcal{S}(\mathbb{R}^d)$  is just the classical space of complex-valued Schwartz functions. A measurable functions  $f : \mathbb{R}^d \rightarrow X$  such that

$$f(\varphi) := \int_{\mathbb{R}^d} dx f(x)\varphi(x) \tag{9}$$

is well-defined for any  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  as a Bochner integral can then be identified via (9) with a distribution in  $\mathcal{S}'(\mathbb{R}^d; X)$ . The elements of  $\mathcal{S}'(\mathbb{R}^d; X)$  can therefore be interpreted as generalized functions with values in  $X$ . Taking  $X = \mathbb{C}$  we simply have

$$\mathcal{S}'(\mathbb{R}^d; \mathbb{C}) = \mathcal{S}'(\mathbb{R}^d),$$

where  $\mathcal{S}'(\mathbb{R}^d)$  is the usual space of tempered distributions.

Most concepts which are known for  $X = \mathbb{C}$  carry over to the general case. We can for instance define addition and multiplication by constants via

$$(f + \tilde{f})(\varphi) := f(\varphi) + \tilde{f}(\varphi), \quad (c \cdot f)(\varphi) := f(c \cdot \varphi).$$

for  $f, \tilde{f} \in \mathcal{S}'(\mathbb{R}^d; X)$ ,  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and  $c \in \mathbb{C}$ . Multiplication with a smooth  $\psi \in C^\infty(\mathbb{R}^d) = C^\infty(\mathbb{R}^d; \mathbb{C})$  whose derivatives grow at most polynomially can further be defined as

$$(\psi \cdot f)(\varphi) := f(\psi \cdot \varphi).$$

The support of a distribution  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  is defined as

$$\text{supp } f := \overline{\{x \in \mathbb{R}^d \mid \forall r > 0 \exists \varphi \in C_c^\infty(B(x, r)) \text{ s.t. } f(\varphi) \neq 0\}}.$$

There is also a Fourier transform  $\mathcal{F}$ ,  $\mathcal{F}^{-1}$  given on  $\mathcal{S}'(\mathbb{R}^d; X)$ , defined via

$$\mathcal{F}f(\varphi) := f(\mathcal{F}\varphi) \quad \text{and} \quad \mathcal{F}^{-1}f(\varphi) := f(\mathcal{F}^{-1}\varphi)$$

for  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  and  $\varphi \in \mathcal{S}(\mathbb{R}^d)$  and where we use the convention

$$\mathcal{F}\varphi(x) = \int_{\mathbb{R}^d} d\xi e^{-2\pi i x \cdot \xi} \varphi(\xi) \quad \text{and} \quad \mathcal{F}^{-1}\varphi(x) = \int_{\mathbb{R}^d} d\xi e^{2\pi i x \cdot \xi} \varphi(\xi),$$

where  $\cdot$  denotes the scalar product on  $\mathbb{R}^d$ . As in the case  $X = \mathbb{C}$  we see that  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  define continuous bijections on  $\mathcal{S}'(\mathbb{R}^d; X)$  that are inverse to each other. As in the scalar case, that is  $X = \mathbb{C}$ , the Fourier transform of every compactly supported  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  is smooth. We further have the usual relations for convolution and multiplication between say  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  and  $\psi \in \mathcal{S}(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d; \mathbb{C})$ :

$$\begin{aligned} \mathcal{F}(\psi * f) &= \mathcal{F}\psi \cdot \mathcal{F}f, & \mathcal{F}^{-1}(\psi * f) &= \mathcal{F}^{-1}\psi \cdot \mathcal{F}^{-1}f, \\ \mathcal{F}(\psi \cdot f) &= \mathcal{F}\psi * \mathcal{F}f, & \mathcal{F}^{-1}(\psi \cdot f) &= \mathcal{F}^{-1}\psi * \mathcal{F}^{-1}f, \end{aligned}$$

where convolutions such as  $\psi * f$  are defined as

$$(\psi * f)(\varphi) := f(\psi * \varphi_{-}).$$

Notice that in contrast to the case  $X = \mathbb{C}$ , in general there is no meaningful convolution concept between two distributions in  $\mathcal{S}'(\mathbb{R}^d; X)$  as  $X$  is not equipped with a product. We use the term *spectral support* to denote the support of the Fourier transform of an  $X$ -valued distribution.

## Besov spaces

The following types of sets will be the building blocks for our definitions below.

**Definition 2.1.** *We say that a set  $\mathfrak{B} \subseteq \mathbb{R}^d$  is a box if there are  $a_1, \dots, a_d > 0$  such that  $\mathfrak{B} = \times_{i=1}^d [-a_i, a_i]$ . A set  $\mathfrak{A} \subseteq \mathbb{R}^d$  is a rectangular annulus if there are two boxes  $\mathfrak{B}, \tilde{\mathfrak{B}} \subseteq \mathbb{R}^d$  with  $\mathfrak{B} \subseteq \tilde{\mathfrak{B}}$  and  $\partial\mathfrak{B} \cap \partial\tilde{\mathfrak{B}} = \emptyset$  such that  $\mathfrak{A} = \tilde{\mathfrak{B}} \setminus \mathfrak{B}$ .*

We then have the following elementary, geometric properties.

**Lemma 2.2.** *Let  $\mathfrak{A}, \tilde{\mathfrak{A}}$  be two rectangular annuli, let  $\mathfrak{B}$  be a box and let  $\mathfrak{s} \in \mathbb{N}_{>}^d$  be a scaling vector. If we define for  $j \geq 0$  the scaled sets  $\mathfrak{A}_j := 2^{j\mathfrak{s}}\mathfrak{A} = \{2^{j\mathfrak{s}}x \mid x \in \mathfrak{A}\}$ ,  $\tilde{\mathfrak{A}}_j := 2^{j\mathfrak{s}}\tilde{\mathfrak{A}}$  and  $\mathfrak{B}_j := 2^{j\mathfrak{s}}\mathfrak{B}$  (with a matrix  $2^{j\mathfrak{s}}$  as in (7)), we have the following relations:*

- $\mathfrak{A}_i \cap \mathfrak{B}_j \neq \emptyset$  only if  $i \lesssim j$ ,
- $\mathfrak{A}_i \cap \tilde{\mathfrak{A}}_j \neq \emptyset$  only if  $i \sim j$ ,

where  $\lesssim$  and  $\sim$  should be read as on page 3.

*Proof.* For the first statement we can write  $2^{i\mathfrak{s}}\mathfrak{A} \cap 2^{j\mathfrak{s}}\mathfrak{B} \neq \emptyset$  as  $\mathfrak{A} \cap 2^{(j-i)\mathfrak{s}}\mathfrak{B} \neq \emptyset$  and use that for  $i \gtrsim j$  this cannot be true by definition of a rectangular annulus. The second statement then follows from the first one if we use that  $\mathfrak{A}$  and  $\tilde{\mathfrak{A}}$  are each contained in some box.  $\square$

We now construct an *anisotropic dyadic partition of unity* for which we essentially follow [Tri06, Section 5.1]. Fix in the following some scaling vector  $\mathfrak{s} \in \mathbb{N}_\times^d$ .<sup>2</sup>

Let  $\mathfrak{B}_{-2} := \emptyset$  and set  $\mathfrak{B}_j := 2^{(j+1)\mathfrak{s}}[-1, 1]^d$  for  $j \geq -1$ . Fix further a symmetric and positive  $\varphi_{-1} \in C_c^\infty(\mathbb{R}^d)$  with values in  $[0, 1]$  such that  $\varphi_{-1} = 1$  on an open set containing  $\mathfrak{B}_{-1} = [-1, 1]^d$  and  $\text{supp } \varphi_{-1} \subseteq \mathfrak{B}_0 = 2^{\mathfrak{s}}[-1, 1]^d$ . We then set for  $j \geq 0$

$$\varphi_j = \varphi_{-1}(2^{-(j+1)\mathfrak{s}} \cdot) - \varphi_{-1}(2^{-j\mathfrak{s}} \cdot),$$

(with *matrices*  $2^{-j\mathfrak{s}}$ ,  $2^{-(j+1)\mathfrak{s}}$  as in (7)). This yields a family  $(\varphi_j)_{j \geq -1} \in C_c^\infty(\mathbb{R}^d)$  that satisfies the following properties:

- $\varphi_j(x) \geq 0$  for  $j \geq -1$  and  $x \in \mathbb{R}^d$ .
- $\sum_{j \geq -1} \varphi_j(x) = 1$  for  $x \in \mathbb{R}^d$ .
- $\text{supp } \varphi_j \subseteq \mathfrak{B}_{j+1} \setminus \mathfrak{B}_{j-1}$  for  $j \geq -1$ , in particular  $\text{supp } \varphi_j \cap \text{supp } \varphi_{j'} = \emptyset$  for  $|j - j'| > 1$ ,

and that further exhibits the following scaling behavior for  $j \geq 0$

$$\varphi_j = \varphi_0(2^{-j\mathfrak{s}} \cdot), \tag{10}$$

$$\sum_{i < j} \varphi_i = \varphi_{-1}(2^{-j\mathfrak{s}} \cdot). \tag{11}$$

The family  $(\varphi_j)_{j \geq -1}$  is called an *(anisotropic) dyadic partition of unity*. Note that for  $j \geq 0$  the support of  $\varphi_j$  is contained in a rectangular annulus of the form  $2^j \mathfrak{A}$  as in Definition 2.1 (namely the set  $\mathfrak{B}_{j+1} \setminus \mathfrak{B}_{j-1}$ ). The support of  $\varphi_{-1}$  is contained in the box  $\mathfrak{B}_0$ .

**Remark 2.3.** *The choice of the sequence  $\mathfrak{B}_j$  was of course rather arbitrary. One could for example have chosen another sequence of boxes such as  $\tilde{\mathfrak{B}}_j = a \cdot \mathfrak{B}_j$  with some  $a > 0$  instead, which would lead to a different partition of unity  $(\tilde{\varphi}_j)_{j \geq -1}$ .*

The following functions will play a special role in this paper:

$$\Psi^j := \mathcal{F}^{-1} \varphi_j, \quad \Psi^{<j} = \sum_{-1 \leq i < j} \Psi^i = \mathcal{F}^{-1} (\varphi_{-1}(2^{-j\mathfrak{s}} \cdot)) \tag{12}$$

with  $j \geq -1$  (note that  $\Psi^{<-1} = 0$ ). We also use occasionally the notation  $\Psi^{\leq j} := \Psi^{<j+1}$  for  $j \geq -1$ .

**Lemma 2.4** (Scaling). *There are real-valued, symmetric Schwartz functions  $\phi_1, \phi_2 \in \mathcal{S}(\mathbb{R}^d)$  such that for  $j \geq 0$*

$$\Psi^j = 2^{j|\mathfrak{s}|} \phi_1(2^{j\mathfrak{s}} \cdot) \quad \text{and} \quad \Psi^{<j-1} = 2^{j|\mathfrak{s}|} \phi_2(2^{j\mathfrak{s}} \cdot),$$

where the matrix  $2^{j\mathfrak{s}}$  should be read as in (7). In particular we have for  $k \in \mathbb{N}^d$

$$\|\partial^k \Psi^j\|_{L^1(\mathbb{R}^d)}, \|\partial^k \Psi^{<j}\|_{L^1(\mathbb{R}^d)} \lesssim 2^{j|k|_{\mathfrak{s}}}.$$

---

<sup>2</sup>To be precise, [Tri06] works with  $\mathfrak{s} \in (0, \infty)^d$ ,  $\sum_{i=1}^d \mathfrak{s}_i = d$  instead, which allows for an interpretation of the Besov regularity in Definition 2.7 below as a sort of *mean* regularity. We here follow the scaling that corresponds to the definitions in [Hai14].

*Proof.* The scaling property follows from (10) and (11) via  $\phi_1 := \mathcal{F}^{-1}\varphi_0$  and  $\phi_2 := \mathcal{F}^{-1}\varphi_{-1}$ . Since  $\varphi_0$  and  $\varphi_{-1}$  are symmetric and real-valued, also their Fourier transforms  $\phi_1$  and  $\phi_2$  are symmetric and real-valued.  $\square$

The action of  $\Psi^{<j}$  on polynomials can be described as follows.

**Lemma 2.5** (Interplay with polynomials). *For  $k, l \in \mathbb{N}^d$  and  $j \geq 0$  we have*

$$\int_{\mathbb{R}^d} du \partial^k \Psi_u^j u^l = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} du \partial^k \Psi_u^{<j} u^l = \delta_{kl} k!.$$

*Proof.* Note that we can replace  $\Psi_u^j$  and  $\Psi_u^{<j}$  in the expressions above by  $\Psi_{-u}^j$  and  $\Psi_{-u}^{<j}$  due to symmetry. We consider the right equality first. By integration by parts the left hand side of this relation equals  $\binom{l}{k} k! \int_{\mathbb{R}^d} du \Psi_{-u}^{<j} u^{l-k}$  (or 0 if  $l > k$ ). This expression can then be rewritten via the inverse Fourier transform as

$$\mathbf{1}_{l \leq k} \binom{l}{k} k! (2\pi i)^{k-l} \partial^{l-k} \mathcal{F}^{-1} \Psi_{-}^{<j}(0),$$

which yields the claim since  $\mathcal{F}^{-1} \Psi^{<j}$  equals 1 in a box around 0. The left relation in the claim is shown in the same way by using  $\mathcal{F}^{-1} \Psi_{-}^j(0) = \varphi_j(0) = 0$  instead.  $\square$

**Lemma 2.6** (Interplay with polynomially growing functions). *If for some  $a \geq 0$  and some measurable  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  we have  $|f_x| \lesssim \|x\|_s^a$ , then this implies for  $k \in \mathbb{N}^d$*

$$\left| \int_{\mathbb{R}^d} du \partial^k \Psi_u^j \cdot f_u \right| + \left| \int_{\mathbb{R}^d} du \partial^k \Psi_u^{<j} \cdot f_u \right| \lesssim 2^{-j(a-|k|_s)}.$$

*Proof.* This follows immediately from Lemma 2.4.  $\square$

Let  $X$  be a Banach space. Using the functions  $\varphi_j$  or their inverse Fourier transforms  $\Psi^j = \mathcal{F}^{-1}\varphi_j$  we define *Littlewood-Paley blocks* for  $X$ -valued distributions  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  and  $j \geq -1$  by

$$\Delta_j f = \mathcal{F}^{-1}(\varphi_j \cdot \mathcal{F}f) = \Psi^j * f = \int_{\mathbb{R}^d} du \Psi_{-u}^j f_u, \quad (13)$$

where multiplication and convolution are defined as in the beginning of this subsection and where we used formal notation in the integral on the right hand side. Note that  $\Delta_j f$  is in  $C^\infty(\mathbb{R}^d; X)$  as the inverse transform of a compactly supported distribution. We can now decompose any  $f \in \mathcal{S}'(\mathbb{R}^d; X)$  by its *Littlewood-Paley decomposition*:

$$f = \sum_{j \geq -1} \Delta_j f, \quad (14)$$

where the sum on the right hand side converges in the topology of  $\mathcal{S}'(\mathbb{R}^d; X)$ . Note that we have for  $p \in [1, \infty]$

$$\|\Delta_j f\|_{L^p(\mathbb{R}^d; X)} = \|\Psi^j * f\|_{L^p(\mathbb{R}^d; X)} \lesssim \|f\|_{L^p(\mathbb{R}^d; X)}, \quad (15)$$

where the involved constant can be chosen independently of  $j$ . Indeed, Young's convolution inequality (which still holds if one of the factors is  $X$ -valued) implies for  $f \in L^p(\mathbb{R}^d; X)$

$$\|\Delta_j f\|_{L^p(\mathbb{R}^d; X)} = \|\Psi^j * f\|_{L^p(\mathbb{R}^d; X)} \leq \|\Psi^j\|_{L^1(\mathbb{R}^d)} \|f\|_{L^p(\mathbb{R}^d; X)} \stackrel{\text{Lemma 2.4}}{\lesssim} \|f\|_{L^p(\mathbb{R}^d; X)},$$

where we used that due to Lemma 2.4 we have  $\|\Psi^j\|_{L^1(\mathbb{R}^d)} = \|\phi_1\|_{L^1(\mathbb{R}^d)} \lesssim 1$  for  $j > -1$ . The same argument shows that

$$S_j f := \sum_{-1 \leq i < j} \Delta_i f = \Psi^{< j} * f = \int_{\mathbb{R}^d} du \Psi_{\cdot - u}^{< j} f_u,$$

is also a bounded operator from  $L^p(\mathbb{R}^d; X)$  to itself for  $p \in [1, \infty]$ . Using the decomposition (14) we can now define anisotropic,  $X$ -valued Besov spaces.

**Definition 2.7.** Let  $\gamma \in \mathbb{R}$ , let  $X$  be a Banach space and let  $\mathfrak{s} \in \mathbb{N}_\times^d$  be a scaling vector. Let further  $(\varphi_j)_{j \geq -1}$  be a dyadic partition of unity on  $\mathbb{R}^d$  defined as above and constructed with  $\mathfrak{s}$ . The anisotropic Besov space  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X)$  is given by

$$\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X) := \left\{ f \in \mathcal{S}'(\mathbb{R}^d; X) \mid \|f\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X)} < \infty \right\},$$

where

$$\|f\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X)} := \left\| \left( 2^{j\gamma} \|\Delta_j f\|_{L^\infty(\mathbb{R}^d; X)} \right)_{j \geq -1} \right\|_{\ell^\infty}. \quad (16)$$

with the Littlewood-Paley blocks  $(\Delta_j)_{j \geq -1}$  defined via  $(\varphi_j)_{j \geq -1}$  as in (13).

**Remark 2.8.** The norm (16) can also be defined with general  $L^p(\mathbb{R}^d; X)$  and  $\ell^q$  norms for  $p, q \in [1, \infty]$  (compare [BCD11, Section 2.7]) which gives rise to a more general space  $\mathcal{B}_{p,q,\mathfrak{s}}^\gamma(\mathbb{R}^d; X)$  that also accounts for the integrability of the considered objects. For the sake of simplicity we will only consider the case  $p = q = \infty$  here.

**Remark 2.9.** Using Lemma 2.2 one sees that another choice of dyadic partition of unity  $(\tilde{\varphi}_j)_{j \geq -1} \subseteq C_c^\infty(\mathbb{R}^d)$  instead of  $(\varphi_j)_{j \geq -1}$ , as in Remark 2.3, gives an equivalent norm for  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X)$ .

We have the following straightforward modification of [BCD11, Lemmas 2.69 and 2.84], see also [GIP15, Lemma A.3].

**Lemma 2.10.** Given a sequence of smooth  $(f_j)_{j \geq -1} \in C^\infty(\mathbb{R}^d; X)$  such that  $\text{supp } \mathcal{F} f_j \subseteq 2^{j\mathfrak{s}} \mathfrak{B}$  for some (fixed) box  $\mathfrak{B}$ , we have for  $\gamma > 0$  and  $f := \sum_{j \geq -1} f_j$

$$\|f\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; X)} \lesssim \left\| \left( 2^{j\gamma} \|f_j\|_{L^p(\mathbb{R}^d; X)} \right)_{j \geq -1} \right\|_{\ell^q}. \quad (17)$$

If  $\text{supp } \mathcal{F} f_j \subseteq 2^{j\mathfrak{s}} \mathfrak{A}$  for a rectangular annulus  $\mathfrak{A}$ , then (17) is even true for all  $\gamma \in \mathbb{R}$ .

An intuition behind the anisotropic scaling is that  $f \in \mathcal{B}_s^\gamma(\mathbb{R}^d; X)$  has “regularity  $\alpha/s_i$  in direction  $i \in \{1, \dots, d\}$ ”<sup>3</sup>. To strenghten this intuition we will find a different characterization of the Besov spaces  $\mathcal{B}_s^\gamma$  based on the Taylor remainder for  $\gamma > 0$

$$R_{x;h}^\gamma f := f - T_{x;h}^\gamma f, \quad (18)$$

where

$$T_{x;h}^\gamma f := \sum_{k \in \mathbb{N}_{<\gamma}^d} \frac{1}{k!} \partial^k f(x) h^k$$

for  $x, h \in \mathbb{R}^d$ ,  $\gamma > 0$ ,  $\mathbb{N}_{<\gamma}^d := \{k \in \mathbb{N}^d \mid |k|_s < \gamma\}$  and with  $f$  having enough derivatives such that these expressions make sense.  $R_{x;h}^\gamma f$  can be rewritten by an application of Proposition 11.1 of [Hai14].

**Lemma 2.11.** *Let  $X$  be a Banach space. Let further  $\gamma \in (0, \infty) \setminus \mathbb{N}$  and let  $f \in C^\infty(\mathbb{R}^d; X)$ . We then have for  $x, h \in \mathbb{R}^d$*

$$R_{x;h}^\gamma f = \sum_{k \in \mathbb{N}_{>\gamma}^d} R_{x;h}^{\gamma,k} f = \sum_{k \in \mathbb{N}_{>\gamma}^d} \frac{h^k}{(k - e_{\mathfrak{m}(k)})!} \int_0^1 dt \partial^k f(x + v_t^k(h)) (1-t)^{k_{\mathfrak{m}(k)}-1}, \quad (19)$$

where  $\mathfrak{m}(k) = \min\{j \mid k_j \neq 0\}$ ,  $\mathbb{N}_{>\gamma}^d := \{k \in \mathbb{N}^d \mid |k|_s > \gamma, |k - e_{\mathfrak{m}(k)}|_s < \gamma\}$  and

$$v_t^k(h) = (h_1, \dots, h_{\mathfrak{m}(k)-1}, t \cdot h_{\mathfrak{m}(k)}, 0, \dots, 0),$$

for the canonical basis  $(e_1, \dots, e_d)$  of  $\mathbb{R}^d$ .

**Remark 2.12.** *The set  $\mathbb{N}_{>\gamma}^d$  can be thought of as the “discrete boundary” of  $\mathbb{N}_{<\gamma}^d$ . Note that this set is finite because it only contains  $k$  with  $|k|_s - \mathfrak{s}_{\mathfrak{m}(k)} < \gamma$ .*

The announced characterization of anisotropic Besov spaces is given by the following lemma, which is a modification of [BCD11, Theorem 2.36].

**Lemma 2.13.** *For  $\gamma \in (0, \infty) \setminus \mathbb{N}$  and a Banach space  $X$  an equivalent norm for  $\mathcal{B}_s^\gamma(\mathbb{R}^d; X)$  is given by the anisotropic Hölder norm*

$$\sup_{l \in \mathbb{N}_{<\gamma}^d} \|\partial^l f\|_{C_b(\mathbb{R}^d; X)} + \sup_{l \in \mathbb{N}_{<\gamma}^d} \sup_{x, y \in \mathbb{R}^d, 0 < \|x-y\|_s \leq 1} \frac{\|\partial^l f(y) - T_{x;y-x}^{\gamma-|l|_s} \partial^l f\|_X}{\|y-x\|_s^{\gamma-|l|_s}}. \quad (20)$$

where we recall that  $\|g\|_{C_b(\mathbb{R}^d; X)} := \sup_{x \in \mathbb{R}^d} \|g(x)\|_X$ .

**Remark 2.14.** *Note that the norm in (20) is equivalent to*

$$\sup_{l \in \mathbb{N}_{<\gamma}^d} \|\partial^l f\|_{C_b(\mathbb{R}^d; X)} + \sup_{l \in \mathbb{N}_{<\gamma}^d} \sup_{x, y \in \mathbb{R}^d, x \neq y} \frac{\|\partial^l f(y) - T_{x;y-x}^{\gamma-|l|_s} \partial^l f\|_X}{\|y-x\|_s^{\gamma-|l|_s}}, \quad (21)$$

since for  $\|x-y\|_s > 1$  the second term of (21) can be bounded via the first term of (20).

<sup>3</sup>Although this intuition is helpful to “guess”  $\mathfrak{s}$  in many situations it is actually slightly incorrect, since the parameter  $\gamma$  should really be read in the sense of an average. A more appropriate (but rather useless) intuition for  $\mathcal{B}_s^\gamma$  would be that  $f \in \mathcal{B}_s^\gamma$  has in average a regularity of  $d \cdot \gamma / |\mathfrak{s}|$ . Compare the regularity of white noise [Hai14, Lemma 10.2] as an example where the “directional intuition” evidently fails.

**Remark 2.15.** *The restriction  $\gamma \notin \mathbb{N}$  is not a shortcoming of our proof: The equivalence of the norms really fails for integers  $\gamma$ , and  $\|\cdot\|_{\mathcal{B}_s^\gamma(\mathbb{R}^d; X)}$  is instead equivalent to a ‘‘Zygmund type’’ norm, see [BCD11, Theorem 2.37] for a result in that direction.*

*Proof.* Assume that  $f \in \mathcal{B}_s^\gamma(\mathbb{R}^d; X)$  as defined in Definition 2.7 above and further, without loss of generality, that  $\|f\|_{\mathcal{B}_s^\gamma(\mathbb{R}^d; X)} \leq 1$ . If we write  $\bar{\Delta}_j f := \sum_{i: |i-j| \leq 1} \Delta_i f$ , we have by spectral support properties  $\Delta_j f = \Delta_j \bar{\Delta}_j f = \Psi^j * \bar{\Delta}_j f$ . Indeed, by construction of  $(\varphi_j)_{j \geq -1}$  we have  $\varphi_j \cdot \sum_{i: |j-i| \leq 1} \varphi_i = \varphi_j \cdot 1 = \varphi_j$  so that by our definition of  $\Delta_j$

$$\mathcal{F}(\Delta_j \bar{\Delta}_j f) = \varphi_j \sum_{i: |j-i| \leq 1} \varphi_i \mathcal{F} f = \varphi_j \mathcal{F} f = \mathcal{F}(\Delta_j f),$$

from which the claimed identity follows. With Lemma 2.4 and Young’s convolution inequality we obtain

$$\begin{aligned} \|\partial^l \Delta_j f\|_{C_b(\mathbb{R}^d; X)} &\stackrel{(*)}{=} \|\partial^l \Delta_j f\|_{L^\infty(\mathbb{R}^d; X)} = \|\partial^l \Psi^j * \bar{\Delta}_j f\|_{L^\infty(\mathbb{R}^d; X)} \\ &\lesssim \|\partial^l \Psi^j\|_{L^1(\mathbb{R}^d)} \|\bar{\Delta}_j f\|_{L^\infty(\mathbb{R}^d; X)} \stackrel{\text{Lem. 2.4}}{\lesssim} 2^{j|l|_s} 2^{-j\gamma} = 2^{-j(\gamma - |l|_s)}, \end{aligned} \quad (22)$$

where we used in  $(*)$  that  $\Delta_j f$  is smooth and in particular continuous. Decomposing  $f = \sum_{j \geq -1} \Delta_j f$  this implies that the first term of (20) is bounded

$$\|\partial^l f\|_{C_b(\mathbb{R}^d; X)} \leq \sum_{j \geq -1} \|\partial^l \Delta_j f\|_{C_b(\mathbb{R}^d; X)} \lesssim \sum_{j > -1} 2^{-j(\gamma - |l|_s)} \lesssim 1.$$

To estimate the second term of (20) we consider for  $j \geq -1$  and  $x, y \in \mathbb{R}^d$  with  $0 < \|x - y\|_s \leq 1$

$$\begin{aligned} (\partial^l \Delta_j f)_y - \sum_{k \in \mathbb{N}_{< \gamma - |l|_s}^d} \frac{(\partial^{k+l} \Delta_j f)_x (y - x)^k}{k!} \\ = \int_{\mathbb{R}^d} du \left( \partial^l \Psi_{y-u}^j - \sum_{k \in \mathbb{N}_{< \gamma - |l|_s}^d} \frac{\partial^{k+l} \Psi_{x-u}^j}{k!} (y - x)^k \right) (\bar{\Delta}_j f)_u, \end{aligned}$$

where we used once more that  $\Delta_j f = \Delta_j \bar{\Delta}_j f = \Psi^j * \bar{\Delta}_j f$  for  $\bar{\Delta}_j f = \sum_{i: |i-j| \leq 1} \Delta_i f$  as above. Formula (19) for the Taylor remainder then gives

$$= \sum_{k \in \mathbb{N}_{> \gamma - |l|_s}^d} \frac{(x - y)^k}{(k - e_m(k))!} \int_0^1 dt \int_{\mathbb{R}^d} du \partial^{k+l} \Psi_{x-u+v_t^k(y-x)}^j (1-t)^{k_m(k)-1} (\bar{\Delta}_j f)_u.$$

With Young’s inequality we thus obtain the bound

$$\begin{aligned} \left\| (\partial^l \Delta_j f)_y - \sum_{k \in \mathbb{N}_{< \gamma - |l|_s}^d} \frac{(\partial^{k+l} \Delta_j f)_x (y - x)^k}{k!} \right\|_X \\ \stackrel{\|y-x\|_s \leq 1}{\lesssim} \sum_{k \in \mathbb{N}_{> \gamma - |l|_s}^d} \|y - x\|_s^{|k|_s} \int_0^1 dt \|(\partial^{k+l} \Psi^j * \bar{\Delta}_j f)_{x+v_t^k(y-x)}\|_X \\ \lesssim \sum_{k \in \mathbb{N}_{> \gamma - |l|_s}^d} \|y - x\|_s^{|k|_s} \int_0^1 dt \|\bar{\Delta}_j f\|_{L^\infty(\mathbb{R}^d; X)} \|\partial^{k+l} \Psi^j\|_{L^1(\mathbb{R}^d)} \\ \lesssim \sum_{k \in \mathbb{N}_{> \gamma - |l|_s}^d} \|x - y\|_s^{|k|_s} 2^{j(|k|_s + |l|_s - \gamma)}, \end{aligned} \quad (23)$$

where we applied  $\|\overline{\Delta}_j f\|_{L^\infty(\mathbb{R}^d; X)} \lesssim 2^{-j\gamma}$  and  $\|\partial^{k+l}\Psi^j\|_{L^1(\mathbb{R}^d)} \lesssim 2^{j(|k|_s+|l|_s)}$  (by Lemma 2.4) in the last step.

On the other hand, by (22), we have the easy estimate

$$\begin{aligned} & \left\| (\partial^l \Delta_j f)_y - \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} \frac{(\partial^{k+l} \Delta_j f)_x}{k!} (y-x)^k \right\|_X \\ & \lesssim 2^{-j(\gamma - |l|_s)} + \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} 2^{-j(\gamma - |l|_s - |k|_s)} \|y-x\|_s^{|k|_s}. \end{aligned} \quad (24)$$

Next, we decompose the Taylor expansion in a “low-frequency” and a “high-frequency” term. That is, choose  $j' \geq -1$  such that  $2^{-j'-1} < \|y-x\|_s \leq 2^{-j'}$  and split

$$\begin{aligned} \partial^l f_y - \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} \frac{(\partial^{k+l} f)_x}{k!} (y-x)^k &= \sum_{j: j \leq j'} [(\partial^l \Delta_j f)_y - \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} \frac{(\partial^{k+l} \Delta_j f)_x}{k!} (y-x)^k] \\ &+ \sum_{j: j > j'} [(\partial^l \Delta_j f)_y - \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} \frac{(\partial^{k+l} \Delta_j f)_x}{k!} (y-x)^k]. \end{aligned}$$

Applying now (23) to the first and (24) to the second term yields the first direction of the equivalence of the norms:

$$\begin{aligned} & \left\| \partial^l f_y - \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} \frac{(\partial^{k+l} f)_x}{k!} (y-x)^k \right\|_X \\ & \lesssim \sum_{j: j \leq j'} \sum_{k \in \mathbb{N}^d_{> \gamma - |l|_s}} \|x-y\|_s^{|k|_s} 2^{j(|k|_s+|l|_s-\gamma)} + \sum_{j: j > j'} \left( 2^{-j(\gamma - |l|_s)} + \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} 2^{-j(\gamma - |l|_s - |k|_s)} \|y-x\|_s^{|k|_s} \right) \\ & \lesssim \sum_{k \in \mathbb{N}^d_{> \gamma - |l|_s}} \|x-y\|_s^{|k|_s} 2^{j'(|k|_s+|l|_s-\gamma)} + \left( 2^{-j'(\gamma - |l|_s)} + \sum_{k \in \mathbb{N}^d_{< \gamma - |l|_s}} 2^{-j'(\gamma - |l|_s - |k|_s)} \|y-x\|_s^{|k|_s} \right) \\ & \lesssim \|y-x\|_s^{\gamma - |l|_s}. \end{aligned}$$

For the opposite direction suppose without loss of generality that (21) is bounded by 1. For  $j > -1$  we then have

$$\begin{aligned} \|(\Delta_j f)_x\|_X &= \left\| \int_{\mathbb{R}^d} du \Psi_{x-u}^j f_u \right\|_X \stackrel{\text{Lem. 2.5}}{=} \left\| \int_{\mathbb{R}^d} du \Psi_{x-u}^j \left( f_u - \sum_{k \in \mathbb{N}^d_{< \gamma}} \frac{(\partial^k f)_x}{k!} (u-x)^k \right) \right\|_X \\ &\leq \int_{\mathbb{R}^d} du |\Psi_{x-u}^j| \cdot \|f_u - T_{x;u-x}^\gamma f\|_X \stackrel{\text{Lem. 2.6}}{\lesssim} 2^{-j\gamma}, \end{aligned}$$

where we used Lemma 2.5 to introduce a term  $\sum_{k \in \mathbb{N}^d_{< \gamma}} \frac{(\partial^k f)_x}{k!} \int_{\mathbb{R}^d} du \Psi_{x-u}^j (u-x)^k = 0$ . To bound  $\Delta_{-1} f$  we apply (15) to get  $\|\Delta_{-1} f\|_{L^\infty(\mathbb{R}^d; X)} \lesssim \|f\|_{L^\infty(\mathbb{R}^d; X)} = \|f\|_{C_b(\mathbb{R}^d; X)} \leq 1$ , which shows that  $\|f\|_{\mathcal{B}_s^\gamma(\mathbb{R}^d; X)} \lesssim 1$ .  $\square$

**Lemma 2.16.** *Let  $X$  be a Banach space, let  $\gamma \in \mathbb{R}$  and  $k \in \mathbb{N}^d$ . We have for  $f \in \mathcal{B}_s^\gamma(\mathbb{R}^d; X)$*

$$\|\partial^k f\|_{\mathcal{B}_s^{\gamma-|k|_s}(\mathbb{R}^d; X)} \lesssim \|f\|_{\mathcal{B}_s^\gamma(\mathbb{R}^d; X)}.$$

*Proof.* As in the beginning of the proof of Lemma 2.13 we use the estimate (22):

$$\|\Delta_j \partial^k f\|_{L^p(\mathbb{R}^d; X)} = \|\partial^k \Delta_j f\|_{L^p(\mathbb{R}^d; X)} \lesssim 2^{j|k|_s} \|\bar{\Delta}_j f\|_{L^p(\mathbb{R}^d; X)}.$$

with  $\bar{\Delta}_j f = \sum_{i: |i-j| \leq 1} \Delta_i f$ , which implies the claim.  $\square$

## 2.2 Basics of regularity structures

In this subsection we recall the basic definition of a regularity structure as a graded vector space equipped with a linear group. We then give a recap on models and on modelled distributions, a concept that is central to this article.

**Definition 2.17.** [Hai14, Definition 2.1] *A regularity structure is a triple  $\mathcal{T} = (A, \mathcal{T}, G)$  consisting of:*

- A locally finite index set  $A \subseteq \mathbb{R}$ , bounded from below, such that  $0 \in A$ .
- A model space  $\mathcal{T} = \bigoplus_{\alpha \in A} \mathcal{T}_\alpha$ , where each  $\mathcal{T}_\alpha$  is a Banach space equipped with a norm  $\|\cdot\|_{\mathcal{T}_\alpha}$ . The space  $\mathcal{T}_0$  is spanned by a unit vector which we call  $\mathbf{1}$ .
- A structure group  $G$  of linear operators acting on  $\mathcal{T}$ , such that for all  $\Gamma \in G$ ,  $\alpha \in A$ ,  $\tau \in \mathcal{T}_\alpha$

$$\Gamma\tau - \tau \in \bigoplus_{\beta < \alpha} \mathcal{T}_\beta. \quad (25)$$

The elements of  $A$  are called homogeneities.

Given  $\tau \in \mathcal{T}$  and  $\alpha \in A$  we write  $\tau^\alpha$  for the projection of  $\tau$  on  $\mathcal{T}_\alpha$ . If  $\dim \mathcal{T}_\alpha < \infty$  and we have a basis  $\{e_i\}$  for  $\mathcal{T}_\alpha$ , we also write  $\tau^{e_j}$  for the coefficient of  $\tau^\alpha$  with respect to  $e_j$ . For example if  $\tau \in \mathcal{T}$  with  $\tau - c \cdot \mathbf{1} \in \bigoplus_{\alpha \in A, \alpha \neq 0} \mathcal{T}_\alpha$  we have  $\tau^\mathbf{1} = c$ . Note that by this definition  $\tau^\alpha \in \mathcal{T}_\alpha$  is a vector in a Banach space, while  $\tau^{e_i}$  is a complex number. We also write

$$\tau^{<\gamma} := \sum_{\alpha \in A: \alpha < \gamma} \tau^\alpha,$$

and similarly for  $\tau^{>\gamma}$ . For  $\Gamma \in G$  we use the abbreviation

$$\Gamma^\alpha \tau := (\Gamma\tau)^\alpha. \quad (26)$$

The same remark applies for the ‘‘basis notation’’ above, for instance  $\Gamma^\mathbf{1} \tau := (\Gamma\tau)^\mathbf{1}$ . We will also need the space

$$\mathcal{T}_\gamma^- := \bigoplus_{\alpha \in A: \alpha < \gamma} \mathcal{T}_\alpha$$

for  $\gamma \in \mathbb{R}$ , so that for example  $\tau^{<\gamma} \in \mathcal{T}_\gamma^-$  for  $\tau \in \mathcal{T}$ . Let us now introduce the notion of a model.

**Definition 2.18.** [Hai14, Definition 2.17] Let  $\mathcal{T} = (A, \mathcal{T}, G)$  be a regularity structure and let  $\mathfrak{s} \in \mathbb{N}_\times^d$  be a scaling vector. A model on  $\mathcal{T}$  with scaling  $\mathfrak{s}$  is a family of linear maps  $\Gamma_{xy} \in G$ ,  $\Pi_x : \mathcal{T} \rightarrow \mathcal{S}'(\mathbb{R}^d)$  for  $x, y \in \mathbb{R}^d$  that satisfy for  $x, y, z \in \mathbb{R}^d$

$$\Gamma_{xx} = \text{Id}_{\mathcal{T}}, \Gamma_{xy}\Gamma_{yz} = \Gamma_{xz}, \Pi_x = \Pi_y\Gamma_{yx}, \quad (27)$$

and further for  $\alpha, \beta \in A$ ,  $\tau \in \mathcal{T}_\alpha$  and  $\beta < \alpha$

$$\|\Gamma_{yx}^\beta \tau\|_{\mathcal{T}_\beta} \lesssim \|\tau\|_{\mathcal{T}_\alpha} \cdot \|x - y\|_{\mathfrak{s}}^{\alpha - \beta}, \quad (28)$$

$$|\Pi_x \tau(\varphi_{\cdot - x}^\lambda)| \lesssim \|\tau\|_{\mathcal{T}_\alpha} \cdot \lambda^\alpha, \quad (29)$$

with  $\varphi_{\cdot - x}^\lambda = \lambda^{-|\mathfrak{s}|} \varphi(\lambda^{-\mathfrak{s}}(\cdot - x))$ , uniformly over all  $\lambda \in (0, 1]$  and  $\varphi \in C_c^\infty(\mathbb{R}^d)$ ,  $\text{supp } \varphi \subseteq B_{\mathfrak{s}}(0, 1)$ ,  $\|\varphi\|_{C^r} \leq 1$ , with  $r \in \mathbb{N}$  being the smallest number strictly bigger than  $-\min A$ . As in (26) we wrote  $\Gamma_{yx}^\beta \tau := (\Gamma_{yx} \tau)^\beta$  for the projection of  $\Gamma_{yx} \tau$  onto  $\mathcal{T}_\beta$ .

We further introduce:

$$\begin{aligned} \|\Pi\|_\gamma &:= \sup_{\varphi} \sup_{\alpha \in A, \alpha < \gamma, \tau \in \mathcal{T}_\alpha, \|\tau\|_{\mathcal{T}_\alpha} \leq 1} \sup_{\lambda \in (0, 1]} \lambda^{-\alpha} |\Pi_x \tau(\varphi_{\cdot - x}^\lambda)|, \\ \|\Gamma\|_\gamma &:= \sup_{x, y \in \mathbb{R}^d, x \neq y} \sup_{\alpha, \beta \in A, \beta < \alpha < \gamma, \tau \in \mathcal{T}_\alpha, \|\tau\|_{\mathcal{T}_\alpha} \leq 1} \|\Gamma_{yx}^\beta \tau\|_{\mathcal{T}_\beta} \|x - y\|_{\mathfrak{s}}^{\beta - \alpha}, \end{aligned}$$

where  $\sup_{\varphi}$  runs over the class of  $\varphi$  described above.

We sometimes write  $\Gamma_{y,x}$  instead of  $\Gamma_{yx}$  to separate the arguments more clearly.

**Remark 2.19.** The functions  $\Pi_x : \mathcal{T} \rightarrow \mathcal{S}'(\mathbb{R}^d)$  do not play an important role in this article and are just mentioned for the sake of completeness. However, compare [GIP15] and [Mar18] for similar concepts to the ones presented here where  $\Pi_x$  becomes important.

Note that we require global bounds on the model  $(\Pi, \Gamma)$  in Definition 2.18, which is different from [Hai14] where the corresponding bounds only need to hold locally uniformly on compact sets. The main reason for requiring global estimates is that we work with an approach based on Fourier analysis, for which it seems unavoidable to work with bounds on the full space. Compare also [HL17] for another work with these assumptions. Global bounds are given immediately in the study of a SPDE if the considered equation is driven by a periodic space white noise. In the case of a spatially periodic space-time white noise we can replace the noise with one that is also periodic in time, with a period that is bigger than the time horizon of the equation. If one wants to consider problems with non-periodic noise, one would have to introduce weights in the analytic bounds of Definition 2.18, similarly as in [HL18, MW17b, MP17]. We will avoid doing so for the sake of simplicity.

A simple example of a regularity structure equipped with a model is the *polynomial regularity structure*  $\overline{\mathcal{T}} = (\overline{A}, \overline{\mathcal{T}}, \overline{G})$ , where

$$\overline{\mathcal{T}} := \text{span} \left\{ X^k \mid k \in \mathbb{N}^d \right\}, \quad (30)$$

where  $\text{span} \{ \dots \}$  denotes the vector space generated by the set in the braces and where we identify  $X^0 = \mathbf{1}$ . We assign the homogeneities  $|X^k| = |k|_{\mathfrak{s}}$  to the symbols  $X^k$  and define  $\overline{A} := \mathbb{N}$ , so that

$$\overline{\mathcal{T}} = \bigoplus_{\alpha \in \overline{A}} \overline{\mathcal{T}}_\alpha := \bigoplus_{\alpha \in \overline{A}} \text{span} \left\{ X^k \mid k \in \mathbb{N}^d, |k|_{\mathfrak{s}} = \alpha \right\}.$$

As each  $\overline{\mathcal{T}}_\alpha = \text{span} \{X^k \mid k \in \mathbb{N}^d, |k|_{\mathfrak{s}} = \alpha\}$  is finite dimensional any choice of norm on  $\overline{\mathcal{T}}_\alpha$  will lead to the same topology; we take

$$\left\| \sum_{k \in \mathbb{N}^d: |k|_{\mathfrak{s}} = \alpha} a_k X^k \right\|_{\overline{\mathcal{T}}_\alpha} := \sum_{k \in \mathbb{N}^d: |k|_{\mathfrak{s}} = \alpha} |a_k|.$$

Consider then the group  $\overline{G} = \{\overline{\Gamma}_h \mid h \in \mathbb{R}^d\}$  with group law  $\overline{\Gamma}_h \overline{\Gamma}_{h'} := \overline{\Gamma}_{h+h'}$  for  $h, h' \in \mathbb{R}^d$  (so that  $\overline{G}$  is essentially  $(\mathbb{R}^d, +)$ ). Fix the action of  $\overline{G}$  on  $\overline{\mathcal{T}}$  by requiring  $\overline{\Gamma}_h X^k := (X + h\mathbf{1})^k$  (with the obvious interpretation of the multiplication on the right hand side).

We can realize a model on  $\overline{\mathcal{F}} = (\overline{A}, \overline{\mathcal{T}}, \overline{G})$  via

$$\overline{\Pi}_x X^k(y) = (y - x)^k, \quad \overline{\Gamma}_{yx} := \overline{\Gamma}_{y-x}, \quad (31)$$

for  $x, y \in \mathbb{R}^d$  and  $k \in \mathbb{N}^d$ . The polynomial regularity structure  $\overline{\mathcal{F}}$  is the example one should have in mind when it comes to comparison of regularity structures with results from “more classical” analysis. From this perspective the spaces  $\mathcal{D}^\gamma$  which we are going to define now (taken from [Hai14, Definition 3.1]) are a generalization of classical Hölder spaces, compare Lemma 2.23 below.

**Definition 2.20.** *Let  $\mathcal{F} = (A, \mathcal{T}, G)$  be a regularity structure, equipped with a model  $(\Pi, \Gamma)$  with scaling  $\mathfrak{s} \in \mathbb{N}_>^d$ . Given  $\gamma \in \mathbb{R}$  we say that a mapping  $F : \mathbb{R}^d \rightarrow \mathcal{T}_\gamma^-$  belongs to  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) = \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}, \Gamma)$  if*

$$\|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})} := \sup_{\alpha \in A, x \in \mathbb{R}^d} \|F_x^\alpha\|_{\mathcal{T}_\alpha} + \sup_{\alpha \in A, x, y \in \mathbb{R}^d, x \neq y} \frac{\|F_y^\alpha - \Gamma_{yx}^\alpha F_x\|_{\mathcal{T}_\alpha}}{\|y - x\|_{\mathfrak{s}}^{\gamma - \alpha}} < \infty. \quad (32)$$

We call the objects in  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  modelled distributions.

**Remark 2.21.** *Due to the bound (28) it is enough to consider in the second term in (32) only pairs  $x, y \in \mathbb{R}^d$  with  $\|x - y\|_{\mathfrak{s}} \leq 1$ , as long as the first supremum in (32) is finite.*

**Remark 2.22.** *Note that all  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  satisfy global bounds, which is different from [Hai14] where the notation  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  only indicates local bounds. In a framework that is largely based on Fourier analysis such as the paracontrolled approach it seems natural to assume global bounds. Local spaces could then be defined afterwards with the help of extension operators, see e.g. Section 5.3.3 of [Mar18] for a version of the Whitney extension theorem in regularity structures.*

The definition of a modelled distribution  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  implies the continuity of every component  $F^\alpha \in C_b(\mathbb{R}^d; \mathcal{T}_\alpha)$  with  $\alpha \in A$  and the bound  $\sup_{\alpha \in A, x \in \mathbb{R}^d} \|F^\alpha(x)\|_{\mathcal{T}_\alpha} < \infty$ . We will denote functions  $F : \mathbb{R}^d \rightarrow \mathcal{T}$  that satisfy these two properties by  $C_b(\mathbb{R}^d; \mathcal{T})$  and set

$$\|F\|_{C_b(\mathbb{R}^d; \mathcal{T})} := \sup_{\alpha \in A, x \in \mathbb{R}^d} \|F_x^\alpha\|_{\mathcal{T}_\alpha},$$

so that in particular  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) \subseteq C_b(\mathbb{R}^d; \mathcal{T})$ . It turns out that on the polynomial regularity structure  $\overline{\mathcal{F}}$  it is rather easy to describe  $\mathcal{D}^\gamma$  in terms of the Besov spaces from Definition 2.7.

**Lemma 2.23.** Let  $\overline{\mathcal{T}} = (\overline{A}, \overline{\mathcal{T}}, \overline{G})$  be the polynomial regularity structure with model  $(\overline{\Pi}, \overline{\Gamma})$  introduced on page 14, for some scaling vector  $\mathfrak{s} \in \mathbb{N}_\times^d$ . Define for  $\gamma \in \mathbb{R}_+ \setminus \mathbb{N}$  the lifted Besov space

$$\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}}) := \{F_f \mid f \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d)\},$$

where  $F_f := \sum_{k \in \mathbb{N}_{<\gamma}^d} \frac{1}{k!} \partial^k f \cdot X^k \in C_b(\mathbb{R}^d; \overline{\mathcal{T}})$  denotes the lift of  $f \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d) = \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathbb{C})$  to the polynomial structure. Equip further  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  with the norm

$$\|F\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})} := \sup_{\alpha \in \overline{A}} \|F^\alpha\|_{\mathcal{B}_\mathfrak{s}^{\gamma-\alpha}(\mathbb{R}^d; \mathcal{T}_\alpha)},$$

where  $F^\alpha$  denotes the projection of  $F \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  on  $\mathcal{T}_\alpha$ . We then have

$$\mathcal{D}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}}) = \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$$

with equivalent norms.

*Proof.* Let  $F \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$ . By definition there is a (unique)  $f \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d)$  such that  $F = F_f$ . In particular we have for  $\alpha \in \overline{A}$

$$F^\alpha = F_f^\alpha = \sum_{k \in \mathbb{N}^d: |k|=\alpha} \frac{1}{k!} \partial^k f \cdot X^k \quad (33)$$

and the norm  $\|F\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})}$  is thus equivalent to

$$\|F\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})} \stackrel{(33)}{\approx} \sup_{k \in \mathbb{N}_{<\gamma}^d} \|\partial^k f\|_{\mathcal{B}_\mathfrak{s}^{\gamma-|k|_\mathfrak{s}}(\mathbb{R}^d)} \stackrel{\text{Lem. 2.16}}{\approx} \|f\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d)}.$$

With Lemma 2.13 we thus get the equivalence

$$\|F\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})} \approx \sup_{l \in \mathbb{N}_{<\gamma}^d} \|\partial^l f\|_{C_b(\mathbb{R}^d)} + \sup_{l \in \mathbb{N}_{<\gamma}^d} \sup_{x, y \in \mathbb{R}^d, 0 < \|x-y\|_\mathfrak{s} \leq 1} \frac{|\partial^l f(y) - T_{x;y-x}^{\gamma-|l|_\mathfrak{s}} \partial^l f|}{\|y-x\|_\mathfrak{s}^{\gamma-|l|_\mathfrak{s}}}. \quad (34)$$

Since  $T_{x;y-x}^{\gamma-|l|_\mathfrak{s}} \partial^l f = \overline{\Gamma}_{yx}^{X^l}(F_f)_x = \overline{\Gamma}_{yx}^{X^l} F_x$  and  $\overline{\Gamma}_{yx}^\alpha F_x = \sum_{l \in \mathbb{N}^d: |l|_\mathfrak{s}=\alpha} \overline{\Gamma}_{yx}^{X^l} F_x \cdot X^l$  for  $\alpha \in \overline{A}$  we see that the right hand side of (34) is equivalent to  $\|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d)}$ , which proves that  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d) \subseteq \mathcal{D}^\gamma(\mathbb{R}^d)$ . The inverse direction follows in the same manner once we show that every  $F \in \mathcal{D}^\gamma(\mathbb{R}^d)$  is of the form  $F = F_f$  for some  $f \in \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d)$ , which can be checked inductively.  $\square$

### 3 Paraproducts on regularity structures

Taking  $X = \mathbb{C}$  in Lemma 2.13 provides two distinct descriptions of the (anisotropic) Besov space  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d) = \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathbb{C})$  with scaling vector  $\mathfrak{s} \in \mathbb{N}_\times^d$  and regularity  $\gamma \in \mathbb{R}_+ \setminus \mathbb{N}$ , given by a Littlewood-Paley characterization in Definition 2.7 and a Hölder norm in Lemma 2.13. In Lemma 2.23 these two characterizations were formulated in the framework of the polynomial regularity structure  $\mathcal{T} = (\overline{A}, \overline{\mathcal{T}}, \overline{G})$  with model  $(\overline{\Pi}, \overline{\Gamma})$ , introduced on page 14. Recall that the statement  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  is just saying that

$$\|F_y^\alpha - \overline{\Gamma}_{yx}^\alpha F_x\|_{\overline{\mathcal{T}}_\alpha} \lesssim \|y-x\|_\mathfrak{s}^{\gamma-\alpha} \quad (35)$$

for  $\alpha \in \overline{A}$ , which can be seen as a Hölder-like characterization similar to the one in Lemma 2.13. On the other hand we introduced in Lemma 2.23 a “Besov space”  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  as the space of functions  $F : \mathbb{R}^d \rightarrow \overline{\mathcal{T}}$  for which it holds

$$\|\Delta_j F^\alpha\|_{L^\infty(\mathbb{R}^d; \overline{\mathcal{T}}_\alpha)} \lesssim 2^{-j(\gamma-\alpha)}. \quad (36)$$

and further

$$F^\alpha = \sum_{|k|_s = \alpha} \frac{\partial^k F^1}{k!} \cdot X^k, \quad (37)$$

which, using the model  $(\overline{\Pi}, \overline{\Gamma})$ , can also be written as

$$\partial^k (F^\alpha - \overline{\Gamma}_x^\alpha F_x)_x = 0 \text{ for } k \in \mathbb{N}_{<\gamma-\alpha}^d, \quad (38)$$

where  $\partial^k (\overline{\Gamma}_x^\alpha F_x)_x$  should be read as the derivative of the map  $y \mapsto \overline{\Gamma}_{yx}^\alpha F_x$  evaluated in the point  $x$ . Note that the relation (37) (or equivalently (38)) can be deduced from (35), but it has to be required explicitly in the definition of  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$  because condition (36) lacks any description of the interplay between different  $F^\alpha, F^{\alpha'}$  with  $\alpha \neq \alpha'$ . However, requiring (38) (or (37)) and (36) we could deduce in Lemma 2.23 that  $\mathcal{D}^\gamma(\mathbb{R}^d; \overline{\mathcal{T}}) = \mathcal{B}_s^\gamma(\mathbb{R}^d; \overline{\mathcal{T}})$ , even without resorting to the explicit construction of  $F_f$  from  $f$ .

Our aim in this article is to find a “Besov space”  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$  on a general regularity structure  $(A, \mathcal{T}, G)$  with model  $(\Pi, \Gamma)$  which describes the space of modelled distributions  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$ . Namely, we want to find a space  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$ , described in terms of Littlewood-Paley blocks, such that

$$\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) = \mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T}). \quad (39)$$

Already in the original paper on paracontrolled distributions [GIP15] the authors introduced a certain paraproduct  $P(F, \Pi)$  on the regularity structure  $(A, \mathcal{T}, G)$  with model  $(\Pi, \Gamma)$  and they conjectured that it might be possible to describe the space  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  via such objects. We here show that this is indeed the case: We introduce a family of paraproducts  $P(F, \Gamma^\alpha)$  and define a space  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$  by requiring instead of (36)

$$\|\Delta_j (F^\alpha - P(F, \Gamma^\alpha))\|_{L^\infty} \lesssim 2^{-j(\gamma-\alpha)}, \quad (40)$$

(which is just saying  $F^\alpha - P(F, \Gamma^\alpha) \in \mathcal{B}_s^{\gamma-\alpha}(\mathbb{R}^d; \mathcal{T}_\alpha)$ ) and the *structure condition* (38) (with  $\overline{\Gamma}$  replaced by  $\Gamma$ ). Since the paraproducts  $P(F, \Gamma^\alpha)$ , described in Definition 3.1 below, vanish for  $F$  with components in the polynomial structure  $\overline{\mathcal{T}}$ , the bound (40) is indeed a generalization of (36).

## Paraproducts

Let us motivate our definitions with a simple example of a singular SPDE, namely the parabolic Anderson model which reads as the following problem on  $[0, \infty) \times \mathbb{R}^2$ :

$$(\partial_t - \Delta_{\mathbb{R}^2})f = f \cdot \xi \quad (41)$$

with periodic space white noise  $\xi \in \mathcal{S}'(\mathbb{R}^2)$  (and a suitable renormalization that we hide for simplicity). The idea in [GIP15] is to define first  $I\xi$  to be the time-independent solution to

$(\partial_t - \Delta_{\mathbb{R}^2})I\xi = -\Delta_{\mathbb{R}^2}I\xi = \xi + \tilde{\xi}$  for an infinitely smooth  $\tilde{\xi}$ , and to consider instead of  $f$  the object

$$f^\# := f - f \otimes I\xi \quad (42)$$

with the paraproduct

$$(f \otimes I\xi)_x = \sum_{j \geq -1} \sum_{i < j-1} (\Delta_i f)_x (\Delta_j I\xi)_x = \sum_{j > 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j f_u \cdot (I\xi)_v, \quad (43)$$

where the integration domain for each integral should be read as *space-time*, that is  $\mathbb{R}^{1+2}$ . Here we cheat a little bit since in [GIP15] a modified paraproduct  $\ll$  in space is considered which allows for a cut-off for negative times. It then turns out that  $f^\#$  solves a “better” equation than  $f$ , which allows to derive a priori estimates and to solve the equation. In (43) we now take functions  $\Psi^j, \Psi^{<j-1}$  that are constructed as in Subsection 2.1 with an *anisotropic* scaling  $\mathfrak{s}$ , more precisely we take in (43) the parabolic scaling  $\mathfrak{s} = \mathfrak{s}_{\text{par}} = (2, 1, 1)$ , which is one more difference with [GIP15].

In [Hai14] the problem (41) is solved on a regularity structure (again with  $\mathfrak{s} = \mathfrak{s}_{\text{par}}$ ), and the solution lives in the subspace spanned by the symbols  $\{\mathcal{I}(\Xi)\} \cup \{X^k : k \in \mathbb{N}^d\}$  and equipped with a model  $(\Pi, \Gamma)$  such that

$$\Pi_x X^k = \Gamma_{yx}^1 X^k = (y - x)^k, \quad \Pi_x \mathcal{I}(\Xi)(y) = \Gamma_{yx}^1 \mathcal{I}(\Xi) = I\xi(y) - I\xi(x). \quad (44)$$

The solution  $f$  to (41) is represented by a modelled distribution  $F$  of the form

$$F = f \mathbf{1} + f \mathcal{I}(\Xi) + \sum_{k \in \mathbb{N}^d: |k|_{\mathfrak{s}_{\text{par}}}=1} f^{X^k} X^k, \quad (45)$$

where  $f^{X^k}$  are some real valued functions and  $f$  is the solution to (41). Recall from Lemma 2.5 that for  $j > 0$  the kernel  $\Psi^j$  integrates polynomials (and constants) to 0, and therefore we can rewrite the paraproduct (43) in terms of  $F$  and the model  $(\Pi, \Gamma)$  as

$$(f \otimes I\xi)_x = \sum_{j > 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j \Gamma_{vu}^1 F_u =: P(F, \Gamma^1). \quad (46)$$

This motivates the following definitions.

**Definition 3.1.** *Let  $\mathcal{T} = (A, \mathcal{T}, G)$  be a regularity structure, let  $(\Pi, \Gamma)$  be a model with scaling  $\mathfrak{s}$  and let  $\Psi^j, \Psi^{<j-1} \in \mathcal{S}(\mathbb{R}^d)$  be functions as in (12) (for the same scaling  $\mathfrak{s}$ ). We define the following paraproducts*

$$P(F, \Gamma^\alpha)_x = \sum_{j > 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j \Gamma_{vu}^\alpha F_u \quad (47)$$

for any  $F : \mathbb{R}^d \rightarrow \mathcal{T}$  and  $\alpha \in A$  for which this is defined. The identity should be read in  $\mathcal{S}'(\mathbb{R}^d; \mathcal{T}_\alpha)$  and is written in formal notation.

**Remark 3.2.** *If  $\mathcal{T}_\alpha$  is finite dimensional for  $\alpha \in A$  and we have a basis  $\{e_i\}$  for  $\mathcal{T}_\alpha$ , we will also write*

$$P(F, \Gamma^{e_i}) := \sum_{j > 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j \Gamma_{vu}^{e_i} F_u,$$

where we recall that  $\Gamma_{vu}^{e_i} F_u$  denotes the coefficient of  $\Gamma_{vu}^\alpha F_u \in \mathcal{T}_\alpha$  in front of  $e_i$ . In particular we have

$$P(F, \Gamma^\alpha) = \sum_i P(F, \Gamma^{e_i}) \cdot e_i.$$

For  $\alpha = 0$  we simply have

$$P(F, \Gamma^0) = P(F, \Gamma^1) \cdot \mathbf{1},$$

a notation we already applied in (46).

**Remark 3.3.** For measurable and at most polynomially growing  $F$  (that is  $\|F^\alpha(x)\|_{\mathcal{T}_\alpha}$  grows at most polynomially in  $x \in \mathbb{R}^d$ ) the expression (47) is well defined. Indeed: Each of the terms in the sum in (47) is spectrally supported in a rectangular annulus of the form  $2^{j^s} \mathfrak{A}$  (where the rectangular annulus  $\mathfrak{A}$  can be chosen independently of  $f$  and  $j$ ). This can be easily checked for smooth  $(x, y) \mapsto \Gamma_{yx} F_x$ , so that the general case follows by approximation. Further by Definition 2.18 and Lemma 2.6 one easily sees that each of the terms can be bounded by  $2^{-j^\kappa}$  for some  $\kappa > 0$ . Lemma 2.10 then shows that  $P(F, \Gamma^\alpha)$  is contained in the Besov space  $\mathcal{B}_s^\kappa(\mathbb{R}^d; \mathcal{T}_\alpha)$  so that  $P(F, \Gamma^\alpha)$  is in fact a (slightly) Hölder continuous function with values in  $\mathcal{T}_\alpha$ .

### The space $\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$

Let us come back to our toy example, the parabolic Anderson model. The function  $f^{X^k}$  from (45) did not appear in the approach of [GIP15], which is due to the fact that according to Lemma 2.5 polynomials are erased in the paraproduct (46). To find a link between the ideas of [GIP15] and [Hai14] we therefore need an extra ingredient that forces the  $f^{X^k}$  to enter the game. This will be the task of the *structure condition*, which we already motivated for the polynomial framework in (38).

**Definition 3.4.** Let  $\mathcal{F} = (A, \mathcal{T}, G)$  be a regularity structure with a model  $(\Pi, \Gamma)$  and scaling  $\mathfrak{s} \in \mathbb{N}^d$ . Construct the functions  $(\Psi^{<N})_{N \geq 0}$ , using the scaling vector  $\mathfrak{s}$ , as in (12). We say that  $F : \mathbb{R}^d \rightarrow \mathcal{T}$  satisfies the structure condition below  $\gamma \in \mathbb{R}$  if for all  $x \in \mathbb{R}^d$ ,  $k \in \mathbb{N}^d$  and  $\alpha \in A$  with  $\alpha < \gamma$  and  $|k|_{\mathfrak{s}} < \gamma - \alpha$  and for all large  $N \in \mathbb{N}$  the map  $v \mapsto \partial^k \Psi_{x-v}^{<N} (F_v^\alpha - \Gamma_{vx}^\alpha F_x)$  is in  $L^1(\mathbb{R}^d; \mathcal{T}_\alpha)$  and it satisfies

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} dv \partial^k \Psi_{x-v}^{<N} (F_v^\alpha - \Gamma_{vx}^\alpha F_x) = 0, \quad (48)$$

where the limit is taken in  $\mathcal{T}_\alpha$ .

**Remark 3.5.** Note that we know from Lemma 2.5 that  $\int_{\mathbb{R}^d} du \Psi_u^{<N} = 1$  so that we see from the scaling property in Lemma 2.4 that  $(\Psi^{<N})_{N \geq 0}$  is a (signed) Dirac sequence. From this fact, or alternatively from the decomposition (14), we conclude that for smooth  $F$  and  $\Gamma$  condition (48) translates into

$$\partial^k (F^\alpha - \Gamma_{\cdot x}^\alpha F_x)_x = 0, \quad (49)$$

which is just the identity we announced in (38). The reason why we do not require (49) from the start is that typically the smoothness assumption is neither satisfied for the modelled distribution  $F$  nor for the map  $y \mapsto \Gamma_{yx} F_x$ , as one may observe in (44).

Let us see how to use the structure condition in order to determine the coefficient  $f^{X^k}$  in (45). Formally using the structure condition in the form (49), and ignoring possible smoothness issues for the sake of simplicity, we obtain for  $\alpha = 0$  and  $k \in \mathbb{N}^d$  with  $|k|_{\mathfrak{s}_{\text{par}}} = 1$  via (45) and (44)

$$0 \stackrel{(49) \text{ for } \alpha = 0}{=} (\partial^k f)_x \mathbf{1} - f_x (\partial^k I\xi)_x \mathbf{1} - k! f_x^{X^k} \mathbf{1},$$

so that the only possible choice for  $f^{X^k}$  is given by

$$f_x^{X^k} = \frac{1}{k!} ((\partial^k f)_x - f_x \cdot (\partial^k I\xi)_x).$$

Although this argument should be executed with more care to guarantee that all expressions are well defined<sup>4</sup>, we can already observe that Definition 3.4 really fixes  $f^{X^k}$  in terms of  $f$ . Equipped with the structure condition there is now hope that we might find a link between the paracontrolled approach in [GIP15] and the description via regularity structures.

As we have already indicated in (42) the object that is really considered in [GIP15] is the difference  $f^\sharp$  of the solution  $f$  of the considered equation with a paraproduct. In view of Definition 3.1 we therefore propose the following definition.

**Definition 3.6.** *Let  $\gamma \in \mathbb{R}$ , let  $\mathcal{T} = (A, \mathcal{T}, G)$  be a regularity structure and let  $(\Pi, \Gamma)$  be a model with scaling  $\mathfrak{s} \in \mathbb{N}^d$ . We say that  $F : \mathbb{R}^d \rightarrow \mathcal{T}_\gamma^-$  with  $F \in C_b(\mathbb{R}^d; \mathcal{T})$  is in  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathcal{T}) = \mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathcal{T}, \Gamma)$  if for all  $\alpha \in A$  with  $\alpha < \gamma$*

$$F^{\sharp, \alpha} := F^\alpha - P(F, \Gamma^\alpha) \in \mathcal{B}_\mathfrak{s}^{\gamma-\alpha}(\mathbb{R}^d; \mathcal{T}_\alpha) \quad (50)$$

and if  $F$  satisfies the structure condition (48) below  $\gamma$ . We define the semi-norm

$$\|F\|_{\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathcal{T})} := \|F\|_{C_b(\mathbb{R}^d; \mathcal{T})} + \sup_{\alpha \in A} \|F^{\sharp, \alpha}\|_{\mathcal{B}_\mathfrak{s}^{\gamma-\alpha}(\mathbb{R}^d; \mathcal{T}_\alpha)}.$$

As we already pointed out above, Definition 3.1 is strictly speaking not a generalization of the approach of [GIP15], because we use space-time paraproduct (43) instead of their modified paraproduct  $\ll$ . The space-time paraproduct might seem more natural, but it comes with a price: Since the solutions to parabolic equations like (41) are only defined for positive times, we have to extend them to negative times to fit them into our framework. We thus need extension operations (and spaces that allow for a blow-up around  $t = 0$ ) in order to practically apply paraproduct techniques to differential equations in regularity structures. We will not deal with these technical issues here and refer to [Mar18] for a few results and concepts in that direction.

## 4 Controlling modelled distributions via paraproducts

We now state and prove the main result of this article. We show that the spaces  $\mathcal{D}^\gamma(\mathbb{R}^d, \mathcal{T})$  from Definition 2.20 and the space  $\mathcal{B}_\mathfrak{s}^\gamma(\mathbb{R}^d; \mathcal{T})$  introduced in Definition 3.6 are identical with equivalent norms. For technical reasons we have to exclude the case that  $\gamma \in \mathbb{R}$  is contained in the locally finite set

$$A_{\mathbb{N}} := A + \mathbb{N}. \quad (51)$$

---

<sup>4</sup>A rigorous approach would consider (48) instead and show the existence of the limit using the paracontrolled structure of  $f$ . Compare the proof of [Mar18, Theorem 6.2.3].

This is necessary since we want to apply for the spaces  $\mathcal{B}_s^{\gamma-\alpha}(\mathbb{R}^d; \mathcal{T}_\alpha)$ , appearing in the Definition of 3.6, the Hölder characterization from Lemma 2.13. We can interpret the following theorem as a generalization of Lemma 2.13, so that the exclusion of  $A_{\mathbb{N}}$  corresponds to the restriction  $\gamma \notin \mathbb{N}$  required there. We do not expect that the result holds for  $\gamma \in A_{\mathbb{N}}$ .

**Theorem 4.1.** *Let  $\mathcal{F} = (A, \mathcal{T}, G)$  be a regularity structure and let  $(\Pi, \Gamma)$  be a model with scaling  $\mathfrak{s} \in \mathbb{N}_x^d$ . We then have for any  $\gamma \in \mathbb{R} \setminus A_{\mathbb{N}}$*

$$\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) = \mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T}) \quad (52)$$

with equivalent norms (where the equivalence constants can both be chosen proportionally to some polynomial in  $\|\Gamma\|_\gamma$ ).

*Proof.* We assume without loss of generality that  $A$  contains only homogeneities below  $\gamma$ . We will include polynomials in  $\|\Gamma\|_\gamma$  in the implicit constant indicated by “ $\lesssim$ ” and we omit the domain “ $\mathbb{R}^d$ ” under integration signs.

For the easy direction  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) \subseteq \mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$  note first that  $F \in \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  already implies that the structure condition (48) is satisfied below  $\gamma$ : Indeed, for  $\alpha \in A$  with  $\alpha < \gamma$  and  $k \in \mathbb{N}_{<\gamma-\alpha}^d$  we have

$$\left\| \int dv \partial^k \Psi_{x-v}^{<N} (F_v^\alpha - \Gamma_{vx}^\alpha F_x) \right\|_{\mathcal{T}_\alpha} \stackrel{\text{Lem. 2.6}}{\lesssim} \|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})} \cdot 2^{-N(\gamma-\alpha-|k|_s)} \xrightarrow{N \rightarrow \infty} 0.$$

To derive the analytic bound we follow similar ideas as in [GIP15, Subsection 6.2]: We can rewrite for  $x \in \mathbb{R}^d$  and  $\alpha \in A$

$$F_x^\alpha - P(F, \Gamma^\alpha)_x = \sum_{j>0} \int dv \left( \Psi_{x-v}^j F_v^\alpha - \int du \Psi_{x-u}^{<j-1} \Psi_{x-v}^j \Gamma_{vu}^\alpha F_u \right) + (\Delta_{\leq 0} F^\alpha)_x.$$

As  $\Delta_{\leq 0} F^\alpha = \Psi^{\leq 0} * F^\alpha$  is smooth with bounded derivatives we only have to consider the first term on the right hand side. We already noted in Remark 3.3 that the  $j$ -th summand is spectrally supported in an annulus  $2^{j\mathfrak{s}}\mathfrak{A}$ , so by Lemma 2.10 it suffices to bound it by  $2^{-j(\gamma-\alpha)} \|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})}$ . We have, using  $\int \Psi_{x-u}^{<j-1} 1 = 1$  (Lemma 2.5),

$$\int dv \left( \Psi_{x-v}^j F_v^\alpha - \int du \Psi_{x-u}^{<j-1} \Psi_{x-v}^j \Gamma_{vu}^\alpha F_u \right) = \iint dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j (F_v^\alpha - \Gamma_{vu}^\alpha F_u).$$

Now, by assumption

$$\|F_v^\alpha - \Gamma_{vu}^\alpha F_u\|_{\mathcal{T}_\alpha} \lesssim \|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})} \|u - v\|_s^{\gamma-\alpha} \lesssim \|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})} (\|u - x\|_s^{\gamma-\alpha} + \|v - x\|_s^{\gamma-\alpha}),$$

so that we have with Lemma 2.6 (and Lemma 2.4) the estimate

$$\left\| \iint dudv \Psi_{x-u}^{<j-1} \Psi_{x-v}^j (F_v^\alpha - \Gamma_{vu}^\alpha F_u) \right\|_{\mathcal{T}_\alpha} \lesssim \|F\|_{\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})} 2^{-j(\gamma-\alpha)},$$

which proves  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T}) \subseteq \mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$ .

Let us now address the delicate direction of the proof, that is  $\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T}) \subseteq \mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$ . Let  $F \in \mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})$  and assume without loss of generality that  $\|F\|_{\mathcal{B}_s^\gamma(\mathbb{R}^d; \mathcal{T})} \leq 1$ . We will show by induction over decreasing homogeneities in  $A$  that for all  $x, y \in \mathbb{R}^d$  and  $\alpha \in A$

$$\|F_y^\alpha - \Gamma_{yx}^\alpha F_x\|_{\mathcal{T}_\alpha} \lesssim \|y - x\|_s^{\gamma-\alpha},$$

which proves the claim. Note that it suffices to take  $\|x - y\|_s \leq 1$ , compare Remark 2.21. We start the induction with  $\alpha = \max A$  (which exists since we assumed  $\max A < \gamma < \infty$ ). By requirement (25) in the definition of  $\Gamma_{yx}$  we have  $\Gamma_{yx}^\alpha F_x = F_x^\alpha$  and thus  $P(F, \Gamma^\alpha) = 0$  due to Lemma 2.5. Hence  $F^\alpha = F_x^{\sharp, \alpha} \in \mathcal{B}_s^{\gamma - \alpha}(\mathbb{R}^d; \mathcal{T}_\alpha)$  and from the structure condition (48) we obtain that  $\partial^k F_x^\alpha = 0$  for  $k \in \mathbb{N}^d$  with  $0 < |k|_s < \gamma - \alpha$  (if any such  $k$  exist). Thus

$$\begin{aligned} \|F_y^\alpha - \Gamma_{yx}^\alpha F_x\|_{\mathcal{T}_\alpha} &= \|F_y^\alpha - F_x^\alpha\|_{\mathcal{T}_\alpha} = \left\| F_y^\alpha - F_x^\alpha - \sum_{k \in \mathbb{N}^d_{< \gamma - \alpha}} \partial^k F_x^\alpha (y - x)^k \right\|_{\mathcal{T}_\alpha} \\ &= \left\| F_y^{\sharp, \alpha} - F_x^{\sharp, \alpha} - \sum_{k \in \mathbb{N}^d_{< \gamma - \alpha}} \partial^k F_x^{\sharp, \alpha} (y - x)^k \right\|_{\mathcal{T}_\alpha} \lesssim \|y - x\|_s^{\gamma - \alpha}, \end{aligned}$$

where we applied Lemma 2.13.

Let us now assume that we already know for some  $\alpha \in A$  that for any  $\alpha' \in A$ ,  $\alpha' > \alpha$

$$\|F_y^{\alpha'} - \Gamma_{yx}^{\alpha'} F_x\|_{\mathcal{T}_{\alpha'}} \lesssim \|y - x\|_s^{\gamma - \alpha'}. \quad (53)$$

We then show that (53) does also hold for all  $\alpha' = \alpha$ . To this end we reshape

$$\begin{aligned} F_y^\alpha - \Gamma_{yx}^\alpha F_x &= F_y^\alpha - F_x^\alpha - \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{yx}^\alpha F_x^{\alpha'} \\ &= F_y^{\sharp, \alpha} - F_x^{\sharp, \alpha} - \sum_{k: 0 < |k|_s < \gamma - \alpha} \frac{1}{k!} \partial^k F_x^{\sharp, \alpha} (y - x)^k \\ &\quad + P(F, \Gamma^\alpha)_y - P(F, \Gamma^\alpha)_x + \sum_{k: 0 < |k|_s < \gamma - \alpha} \frac{1}{k!} \partial^k (F^\alpha - P(F, \Gamma^\alpha))_x (y - x)^k - \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{yx}^\alpha F_x^{\alpha'}. \end{aligned} \quad (54)$$

Since (54) already decays in the right order due to Lemma 2.13 (and the assumption  $\gamma \notin A_{\mathbb{N}}$ ), we are only left with the last line which below we identify as the limit for  $N \rightarrow \infty$  (in  $\mathcal{T}_\alpha$  for every  $x, y$ ) of

$$\begin{aligned} D^N &:= \sum_{j \leq N} D_j^N \\ &:= \sum_{j \leq N} \left[ \int dw \left( \Psi_{y-w}^j - \sum_{k \in \mathbb{N}^d_{< \gamma - \alpha}} \frac{1}{k!} \partial^k \Psi_{x-w}^j (y - x)^k \right) \left( P(F, \Gamma^\alpha)_w - \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{wx}^\alpha F_x^{\alpha'} \right) \right]. \end{aligned}$$

Indeed, we have the following three convergences:

$$\sum_{j \leq N} \int dw \left( \Psi_{y-w}^j - \Psi_{x-w}^j \right) P(F, \Gamma^\alpha)_w \quad (55)$$

$$\begin{aligned} &\xrightarrow{N \rightarrow \infty} P(F, \Gamma^\alpha)_y - P(F, \Gamma^\alpha)_x, \\ &- \sum_{j \leq N} \sum_{0 < |k|_s < \gamma - \alpha} \int dw \frac{1}{k!} \partial^k \Psi_{x-w}^j \left( P(F, \Gamma^\alpha)_w - \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{wx}^\alpha F_x^{\alpha'} \right) (y - x)^k \end{aligned} \quad (56)$$

$$\begin{aligned} &\xrightarrow{N \rightarrow \infty} \sum_{0 < |k|_s < \gamma - \alpha} \frac{1}{k!} \partial^k (F^\alpha - P(F, \Gamma^\alpha))_x (y - x)^k, \\ &\sum_{j \leq N} \int dw \left( \Psi_{y-w}^j - \Psi_{x-w}^j \right) \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{wx}^\alpha F_x^{\alpha'} \end{aligned} \quad (57)$$

$$\xrightarrow{N \rightarrow \infty} \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{yx}^\alpha F_x^{\alpha'} - \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{xx}^\alpha F_x^{\alpha'} = \sum_{\alpha' \in A: \alpha' > \alpha} \Gamma_{yx}^\alpha F_x^{\alpha'},$$

where we used (as in Remark 3.5) that  $\Psi^{\leq N} = \sum_{j \leq N} \Psi^j$  is a Dirac sequence, for the second term in (56) we applied the structure condition (48), and in (57) we used and the continuity of  $\Gamma_{\cdot x}^\alpha F_x^{\alpha'}$ <sup>5</sup>. Writing  $D^N = (55) + (56) + (57)$  we see the claimed convergence of  $D^N$ . We now show that, uniformly in  $N$ ,

$$\|D^N\|_{\mathcal{T}_\alpha} \lesssim \|y - x\|_s^{\gamma - \alpha}. \quad (58)$$

Note that we can reshape (with  $\Delta_{>0} := \sum_{j>0} \Delta_j$ )

$$\begin{aligned} P(F, \Gamma^\alpha)_w &= \sum_{j>0} \iint dudv \Psi_{w-u}^{<j-1} \Psi_{w-v}^j \Gamma_{vu}^\alpha F_u \\ &= \sum_{j>0} \iint dudv \Psi_{w-u}^{<j-1} \Psi_{w-v}^j \Gamma_{vu}^\alpha (F_u - \Gamma_{ux}^\alpha F_x) + (\Delta_{>0} \Gamma_{\cdot x}^\alpha F_x)_w \\ &= \sum_{j>0} \sum_{\alpha' \in A: \alpha' > \alpha} \iint dudv \Psi_{w-u}^{<j-1} \Psi_{w-v}^j \Gamma_{vu}^\alpha (F_u^{\alpha'} - \Gamma_{ux}^{\alpha'} F_x) + \sum_{\alpha' \in A: \alpha' > \alpha} (\Delta_{>0} \Gamma_{\cdot x}^\alpha F_x^{\alpha'})_w, \end{aligned}$$

where in the last line we used (25) for both terms and also that  $\int dv \Psi_v^j 1 = 0$  for  $j > 0$  (by Lemma 2.5) to cancel the  $\alpha' = \alpha$  components. We can therefore rewrite  $D_j^N$  as follows (with  $R_{x;y-x}^{\gamma-\alpha}$  being the Taylor remainder as in (18))

$$\begin{aligned} D_j^N &= \sum_{j \leq N} \sum_{\alpha' \in A: \alpha' > \alpha} \int dw R_{x-w;y-x}^{\gamma-\alpha} \Psi^j \\ &\quad \times \left[ \sum_{i>0} \iint dudv \Psi_{w-u}^{<i-1} \Psi_{w-v}^i \Gamma_{vu}^\alpha (F_u^{\alpha'} - \Gamma_{ux}^{\alpha'} F_x) + (\Delta_{\leq 0} \Gamma_{\cdot x}^\alpha F_x^{\alpha'})_w \right] \\ &\stackrel{(*)}{=} \sum_{j \leq N} \sum_{\alpha' \in A: \alpha' > \alpha} \int dw R_{x-w;y-x}^{\gamma-\alpha} \Psi^j \left[ \sum_{i>0: i \sim j} \iint dudv \Psi_{w-u}^{<i-1} \Psi_{w-v}^i \Gamma_{vu}^\alpha (F_u^{\alpha'} - \Gamma_{ux}^{\alpha'} F_x) \right] \\ &\quad + \sum_{\alpha' \in A: \alpha' > \alpha} \int dw \Psi_w^{<N+1} \cdot R_{x-w;y-x}^{\gamma-\alpha} (\Delta_{\leq 0} \Gamma_{\cdot x}^\alpha F_x^{\alpha'}), \end{aligned} \quad (60)$$

where in (\*) we used spectral support properties to restrict the inner sum to  $i \sim j$  and the convolution-like structure to move the Taylor remainder to  $\Delta_{\leq 0} \Gamma_{\cdot x}^\alpha F_x^{\alpha'}$  in the last term. The last term can be estimated by  $\|x - y\|_s^{\gamma - \alpha}$  via Lemma 2.11 and Lemma 2.6 if we use that for every  $k \in \mathbb{N}^d$  there is a  $C > 0$  such that

$$\|\partial^k \left( \Delta_{\leq 0} \Gamma_{\cdot x}^\alpha F_x^{\alpha'} \right)_{w-x+v_t^k(y-x)}\|_{\mathcal{T}_\alpha} \lesssim (1 + \|w\|_s^C), \quad (61)$$

which can be easily checked by direct computation. To handle the term (60) we first estimate the sum in the square brackets: By Definition 2.18 and the induction hypothesis we have that

$$\sum_{\alpha' \in A: \alpha' > \alpha} \|\Gamma_{vu}^\alpha (F_u^{\alpha'} - \Gamma_{ux}^{\alpha'} F_x)\|_{\mathcal{T}_\alpha} \lesssim \sum_{\alpha' \in A: \alpha' > \alpha} \|v - u\|_s^{\alpha' - \alpha} \cdot \|u - x\|_s^{\gamma - \alpha'}.$$

<sup>5</sup>A short computation shows that Definition 2.18 already implies (Hölder) continuity of the maps  $y \mapsto \Gamma_{yx}^\alpha \tau$  for  $\tau \in \mathcal{T}$  and  $x \in \mathbb{R}^d$ .

Lemma 2.6 thus yields

$$\begin{aligned} \sum_{i: i \sim j} \sum_{\alpha' \in A: \alpha' > \alpha} \left\| \iint dudv \Psi_{w-v}^{<i-1} \Psi_{w-u}^i \Gamma_{vu}^\alpha (F_u^{\alpha'} - \Gamma_{ux}^{\alpha'} F_x) \right\|_{\mathcal{T}_\alpha} \\ \lesssim \sum_{\alpha' \in A: \alpha' > \alpha} 2^{-j(\alpha' - \alpha)} 2^{-j(\gamma - \alpha')} \lesssim 2^{-j(\gamma - \alpha)}. \end{aligned} \quad (62)$$

The rest of the estimate for (60) then follows by the same line of reasoning as in Lemma 2.13: Let  $j'$  be such that  $2^{-j'-1} < \|x - y\|_s \leq 2^{-j'}$  and bound the sum in (60) (up to a constant factor) using (62) by

$$\sum_{j \leq j'} \sum_{k \in \mathbb{N}_{>\gamma-\alpha}^d} \|x - y\|_s^{|k|_s} 2^{j(|k|_s - (\gamma - \alpha))} + \sum_{N \geq j > j'} \sum_{k \in \mathbb{N}_{<\gamma-\alpha}^d} \|x - y\|_s^{|k|_s} 2^{-j(\gamma - \alpha - |k|_s)} \lesssim \|x - y\|_s^{\gamma - \alpha},$$

where we applied Lemma 2.11 in the low-frequency case and in both cases the  $L^1$ -bound from Lemma 2.4. We have thus shown (58) so that in total

$$\|F_y^\alpha - \Gamma_{yx}^\alpha F_x\|_{\mathcal{T}_\alpha} = \lim_{N \rightarrow \infty} \|D_N\|_{\mathcal{T}_\alpha} \lesssim \|x - y\|_s^{\gamma - \alpha},$$

which closes the induction and finishes the proof.  $\square$

**Remark 4.2.** In [HL17] the authors introduce a more general space  $\mathcal{D}_{p,q}^\gamma(\mathbb{R}^d; \mathcal{T})$  which generalizes the space  $\mathcal{D}^\gamma(\mathbb{R}^d; \mathcal{T})$  from Definition 2.20 (the special case  $\mathcal{D}_{p,p}^\gamma(\mathbb{R}^d)$  was already introduced before by [PT16]). Our approach would allow us to also define a ‘‘Besov space’’  $\mathcal{B}_{p,q,s}^\gamma(\mathbb{R}^d; \mathcal{T})$  (compare Remark 2.8), and we expect that it is equal to  $\mathcal{D}_{p,q}^\gamma(\mathbb{R}^d)$ . But for the sake of simplicity we restrict ourselves to the case  $p = q = \infty$  in this work.

## References

- [AC15] R. Allez and K. Chouk. The continuous Anderson hamiltonian in dimension two. *ArXiv e-prints*, November 2015.
- [BCD11] H. Bahouri, J.-Y. Chemin, and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations*. Springer-Verlag, 2011.
- [CC18a] Giuseppe Cannizzaro and Khalil Chouk. Multidimensional sdes with singular drift and universal construction of the polymer measure with white noise potential. *Ann. Probab.*, 46(3):1710–1763, 05 2018.
- [CC18b] Rémi Catellier and Khalil Chouk. Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *The Annals of Probability*, 2018.
- [CFG17] Giuseppe Cannizzaro, Peter K Friz, and Paul Gassiat. Malliavin calculus for regularity structures: The case of gPAM. *J. Funct. Anal.*, 272(1):363–419, 2017.
- [FH14] Peter K. Friz and Martin Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Springer International Publishing, Cham, 2014.
- [GIP15] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3, 2015.
- [GP17] Massimiliano Gubinelli and Nicolas Perkowski. KPZ reloaded. *Communications in Mathematical Physics*, 349(1):165–269, Jan 2017.
- [Hai14] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.
- [HL17] Martin Hairer and Cyril Labbé. The reconstruction theorem in Besov spaces. *Journal of Functional Analysis*, 273(8):2578 – 2618, 2017.

- [HL18] Martin Hairer and C. Labbe. Multiplicative stochastic heat equations on the whole space. *Journal of the European Mathematical Society*, June 2018.
- [HM18] M. Hairer and K. Matetski. Discretisations of rough stochastic pdes. *Ann. Probab.*, 46(3):1651–1709, 05 2018.
- [HP15] Martin Hairer and Étienne Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan*, 67(4):1551–1604, 2015.
- [HQ15] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ. *ArXiv e-prints*, December 2015.
- [HS17] Martin Hairer and Hao Shen. A central limit theorem for the kpz equation. *Ann. Probab.*, 45(6B):4167–4221, 11 2017.
- [KM17] A. Kupiainen and M. Marozzi. Renormalization of Generalized KPZ Equation. *Journal of Statistical Physics*, 166:876–902, 2017.
- [Kup16] Antti Kupiainen. Renormalization group and stochastic pdes. *Annales Henri Poincaré*, 17(3):497–535, 2016.
- [Mar18] J. Martin. *Refinements of the Solution Theory for Singular SPDEs*. Dissertation, Humboldt-Universität zu Berlin, 2018.
- [MP17] J. Martin and N. Perkowski. Paracontrolled distributions on Bravais lattices and weak universality of the 2d parabolic Anderson model. *ArXiv e-prints*, April 2017.
- [MW17a] J.-C. Mourrat and H. Weber. The Dynamic  $\Phi_3^4$  Model Comes Down from Infinity. *Communications in Mathematical Physics*, 356:673–753, December 2017.
- [MW17b] Jean-Christophe Mourrat and Hendrik Weber. Global well-posedness of the dynamic  $\phi^4$  model in the plane. *Ann. Probab.*, 45(4):2398–2476, 07 2017.
- [OSSW18] Felix Otto, Jonas Sauer, Scott Smith, and Hendrik Weber. Parabolic equations with rough coefficients and singular forcing. *arXiv preprint arXiv:1803.07884*, 2018.
- [OW16] Felix Otto and Hendrik Weber. Quasilinear SPDEs via rough paths. *arXiv preprint arXiv:1605.09744*, 2016.
- [PT16] D. J. Prömel and J. Teichmann. Stochastic Analysis with Modelled Distributions. *ArXiv e-prints*, September 2016.
- [Sch50] Laurent Schwartz. *Théorie des distributions*, volume 1. Publications de l’Institut de Mathématique de l’Université de Strasbourg, 1950.
- [ST87] H.J. Schmeisser and H. Triebel. *Topics in Fourier analysis and function spaces*. Wiley, 1987.
- [Tre75] François Trèves. 39 functions and distributions valued in banach spaces. In *Basic Linear Partial Differential Equations*, volume 62 of *Pure and Applied Mathematics*, pages 381 – 390. Elsevier, 1975.
- [Tri06] Hans Triebel. *Theory of Function Spaces III*. Birkhäuser Basel, Basel, 2006.
- [ZZ18] Rongchan Zhu and Xiangchan Zhu. Lattice approximation to the dynamical  $\phi_3^4$  model. *Ann. Probab.*, 46(1):397–455, 01 2018.