



The Wave Maps Equation and Brownian Paths

Bjoern Bringmann^{1,2} , Jonas Lührmann³, Gigliola Staffilani⁴

¹ School of Mathematics, Institute for Advanced Study, Princeton, NJ 08540, USA. E-mail: bjoern@ias.edu

² Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

³ Department of Mathematics, Texas A&M University, College Station, TX 77843, USA.

E-mail: luehrmann@math.tamu.edu

⁴ Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

E-mail: gigliola@math.mit.edu

Received: 30 September 2022 / Accepted: 9 November 2023

Published online: 23 February 2024 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract: We discuss the $(1+1)$ -dimensional wave maps equation with values in a compact Riemannian manifold \mathcal{M} . Motivated by the Gibbs measure problem, we consider Brownian paths on the manifold \mathcal{M} as initial data. Our main theorem is the probabilistic local well-posedness of the associated initial value problem. The analysis in this setting combines analytic, geometric, and probabilistic methods.

Contents

1. Introduction	2
1.1 Motivation	2
1.2 Related works	4
1.2.1 Random dispersive equations	4
1.2.2 The wave maps equation	6
1.2.3 Singular parabolic SPDEs	7
1.3 Main result and proof ideas	8
1.3.1 Low \times high-interactions	10
1.3.2 High \times high \rightarrow low-interactions	12
1.3.3 Further remarks	13
1.4 Open problem: invariant Gibbs measure	14
2. Preparations	15
2.1 Notation and parameters	15
2.2 Function spaces, para-products and product estimates	17
2.3 Commutator estimates	23
2.4 Integrals and traces	27
2.5 Proof of Theorem 1.2.(i)	31
2.6 Brownian paths	34
3. Ansatz	48
3.1 The shifted wave map	48

3.2	The enhanced initial data	50
3.3	Rigorous definition of the Ansatz	52
3.4	The modulation and forced wave maps equations	55
4.	Prototypical Estimates	57
4.1	Favorable derivatives	57
4.2	Prototypical bilinear and trilinear estimates	59
4.3	Unfortunate resonances	61
4.4	Generalizations of (+) and (+−)	62
5.	Bilinear Analysis	63
5.1	The null-form $\partial_u \phi^i \partial_v \phi^j$	63
5.2	The product $\phi^k \partial_u \phi^i$	70
6.	The Full Nonlinearity	74
6.1	The main term	74
6.1.1	The case $N \leq M^{1-\delta}$	75
6.1.2	The case $M \sim_\delta N$	77
6.2	The para-controlled and resonance terms	82
6.3	The remainder terms	85
6.4	Proof of Theorem 6.1	85
7.	Duhamel Integral Approximation	86
7.1	The low \times high-term	87
7.2	The high \times high-term	91
7.3	Proof of Proposition 7.1	93
8.	Modulation Equations	93
8.1	Para-controlled modulation equations	94
8.2	Estimates for the modulation equations	98
8.3	Proof of Proposition 3.14	101
9.	Local Well-Posedness	104
A.	Appendix A. Deterministic Ill-Posedness	108

1. Introduction

1.1. Motivation. The wave maps equation is a geometric generalization of the linear wave equation for scalar-valued fields to fields that take values in a Riemannian manifold. Formally, a wave map $\phi: \mathbb{R}^{1+1} \rightarrow \mathcal{M}$ from $(1+1)$ -dimensional Minkowski space \mathbb{R}^{1+1} into a Riemannian manifold (\mathcal{M}, g) is a critical point of the Lagrangian action functional

$$\mathcal{L}(\phi) := \iint_{\mathbb{R}^{1+1}} (-|\partial_t \phi|_g^2 + |\partial_x \phi|_g^2) dx dt.$$

Throughout this paper we work with the extrinsic formulation of the wave maps equation and we consider smooth, compact Riemannian target manifolds (\mathcal{M}, g) without boundary. By Nash's embedding theorem, we may regard \mathcal{M} as an isometrically embedded submanifold of some Euclidean space \mathbb{R}^D . We write $\phi = (\phi^1, \dots, \phi^D): \mathbb{R}^{1+1} \rightarrow \mathcal{M} \hookrightarrow \mathbb{R}^D$ for the corresponding extrinsic representation of the wave map and we denote by $\mathbb{I}: T\mathcal{M} \times T\mathcal{M} \rightarrow N\mathcal{M}$ the second fundamental form of the embedding $\mathcal{M} \hookrightarrow \mathbb{R}^D$. Then the wave maps equation from $(1+1)$ -dimensional Minkowski space to \mathcal{M} takes the form

$$\partial_\mu \partial^\mu \phi^k = -\mathbb{I}_{ij}^k(\phi) \partial_\mu \phi^i \partial^\mu \phi^j, \quad 1 \leq i, j, k \leq D, \quad 0 \leq \mu \leq 1. \quad (\text{WM})$$

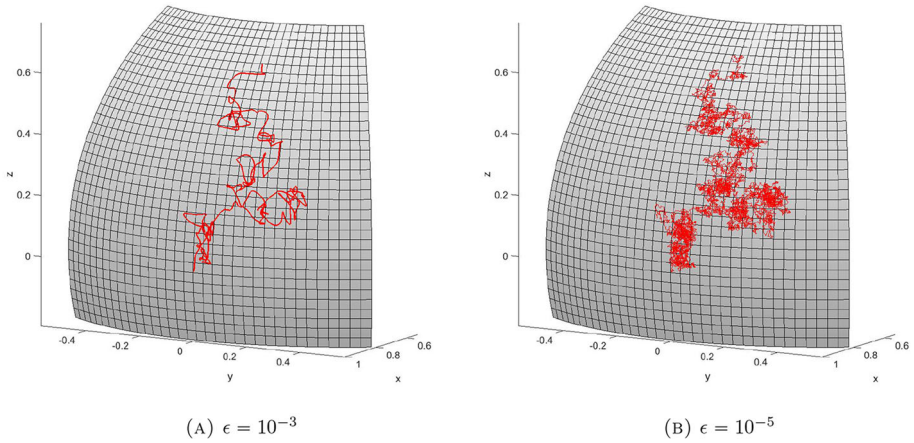


Fig. 1. We display the approximations B^ϵ of the Brownian path B on \mathbb{S}^2 for two different values of ϵ . For $\epsilon = 10^{-3}$, the path still looks relatively smooth. For $\epsilon = 10^{-5}$, the path already exhibits heavy oscillations on small scales. In spite of these oscillations, the behavior of the two approximation agrees on large scales

We use the standard conventions of raising or lowering indices with respect to the Minkowski metric with signature $(-+)$ on \mathbb{R}^{1+1} , and of summing over repeated upper and lower indices.

The equation (WM) is invariant under the scaling

$$\phi(t, x) \mapsto \phi(\lambda t, \lambda x), \quad \lambda > 0, \quad (1.1)$$

and solutions formally conserve the Hamiltonian

$$H[\phi, \phi_t] := \frac{1}{2} \int_{\mathbb{R}} (|\partial_x \phi|^2 + |\partial_t \phi|^2) dx. \quad (1.2)$$

Our interest lies in the Cauchy problem for (WM) for initial data

$$(\phi, \partial_t \phi)|_{t=0} = (\phi_0, \phi_1)$$

given by maps

$$\phi_0: \mathbb{R} \rightarrow \mathcal{M}, \quad \phi_1: \mathbb{R} \rightarrow \phi_0^* T\mathcal{M}, \quad \phi_1(x) \in T_{\phi_0(x)} \mathcal{M}. \quad (1.3)$$

Motivated by the problem of constructing and proving the invariance of the Gibbs measure for the wave maps equation, we choose the initial data ϕ_0 as a Brownian path on the manifold and the initial velocity ϕ_1 as white noise on the pullback bundle $\phi_0^* T\mathcal{M}$. A Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ is the natural generalization of Euclidean Brownian motion $W: \mathbb{R} \rightarrow \mathbb{R}^D$ from the vector-valued to the manifold-valued case. As we will describe below (Sect. 1.3 and Sect. 2.6), the Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ can be approximated by smooth paths $(B^\epsilon)_\epsilon: \mathbb{R} \rightarrow \mathcal{M}$, which are illustrated in Fig. 1. In the literature, Brownian paths are more commonly referred to as Brownian motion on a manifold, but our terminology better distinguishes the argument of $B = B(x)$ from the time-coordinate in the wave maps equation (WM). Brownian paths are natural geometric objects and have important applications to heat kernel estimates and index theorems. For a detailed introduction, we refer the reader to the textbook [Hsu02]. The initial

velocity $V \in B^*T\mathcal{M}$, which is described in more detail below, is chosen as the natural generalization of vector-valued white noise.

The Gibbs measure of the wave maps equation is formally given by

$$d\mu(\phi_0, \phi_1) = \mathcal{Z}^{-1} \exp \left(-\frac{1}{2} \int_{\mathbb{R}} |\partial_x \phi_0|_g^2 + |\phi_1|_g^2 dx \right) d\phi_0 d\phi_1, \quad (1.4)$$

where \mathcal{Z} is a normalization constant. While the Gibbs measure in (1.4) has not yet been constructed rigorously, we believe that the pair (B, V) accurately describes its samples. This belief is based on the Wiener measure, which is the analogue of the Gibbs measure in the parabolic setting (cf. [Hai16, BGHZ21]). The Wiener measure is formally given by

$$d\nu(\phi_0) = \mathcal{Z}^{-1} \exp \left(-\frac{1}{2} \int_{\mathbb{R}} |\partial_x \phi_0|_g^2 dx \right) d\phi_0$$

and has been constructed¹ rigorously by Andersson and Driver in [AD99]. Furthermore, [AD99] proves that the Wiener measure is absolutely continuous with respect to the law of Brownian paths, where the Radon-Nikodym derivative only depends on the scalar curvature of the manifold.

We now state an informal version of our main result.

Theorem 1.1 (Informal version). *The wave maps equation (WM) with initial data $(\phi_0, \phi_1) = (B, V)$ is almost surely locally well-posed.*

The statement and proof of Theorem 1.1 involve analytic, geometric, and probabilistic aspects, which is our motivation for proving this theorem. As mentioned above and further detailed below (see Sect. 1.4), Theorem 1.1 can also be seen as partial progress towards proving the invariance of the Gibbs measure for the wave maps equation.

1.2. Related works. The study of the $(1+1)$ -dimensional wave maps evolution of Brownian paths on Riemannian manifolds features intriguing connections to several research areas in partial differential equations. In this section we describe some of these connections and we highlight related results on the study of the initial value problem for dispersive and hyperbolic equations with random initial data, on the deterministic well-posedness of (semi-linear) geometric wave equations, and on singular parabolic stochastic PDEs. In view of the rich and vast literature on these subjects, we do not attempt to be exhaustive.

1.2.1. Random dispersive equations In recent years, there has been tremendous interest in random dispersive equations. In this introduction, we only discuss selected works in this field and refer the reader to the surveys [BOP19, NS19], the introduction of [Bri20], and the original works [BOP15, Bou94, Bou96, Bri18, Bri20, BDNY22, BT08a, BT08b, CG19, DH19, DH21, DLM20, DNY19, DNY20, DNY21, GKO18, KLS20, KM19, KMV20, LM14, NORBS12, OOT20, OOT21, OST21, ST21, Tzv15].

The main motivation of this work is the Gibbs measure problem for Hamiltonian PDEs. At a formal level, we consider a symplectic manifold (\mathcal{N}, ω) of dimension

¹ To be precise, [AD99] constructed the Wiener measure on the compact domain $[0, 1]$ instead of \mathbb{R} .

$\dim(\mathcal{N}) = 2n$, where $n \in \mathbb{N} \cup \{\infty\}$, and a Hamiltonian $H: \mathcal{N} \rightarrow \mathbb{R}$. Then, the Hamiltonian equation and Gibbs measure are formally given by

$$\dot{\phi} = \nabla_{\omega} H(\phi) \quad (1.5)$$

and

$$\mu = \mathcal{Z}^{-1} \exp(-H) \omega^n, \quad (1.6)$$

where \mathcal{Z} is a normalization constant and ω^n is the n -fold wedge product of ω .

Gibbs measure problem. *Prove the existence of the Gibbs measure and its invariance under the Hamiltonian flow.*

In the following, we only discuss proofs of invariance, and refer to aspects regarding the existence of the Gibbs measure to [ADC21, GJ87, GH21]. To fix ideas, we further restrict our discussion to the (renormalized) periodic defocusing nonlinear Schrödinger equation (NLS). In dimension $d = 2$, it is given by²

$$i\partial_t \phi + \Delta \phi =: |\phi|^{p-1} \phi: \quad (t, x) \in \mathbb{R} \times \mathbb{T}^2. \quad (1.7)$$

Here, $:|\phi|^{p-1} \phi:$ denotes the Wick-ordering of $|\phi|^{p-1} \phi$. Inspired by Lebowitz-Rose-Speer [LRS88], the seminal work of Bourgain [Bou94] solved the Gibbs measure problem for the (NLS) in dimension $d = 1$. In later work [Bou96], Bourgain also solved this problem for (NLS) with $(d, p) = (2, 3)$, i.e., the cubic nonlinear Schrödinger equation in two dimensions. The main difficulties in [Bou96] stem from the low regularity of the initial data. Let ϕ_0 be a random sample from the corresponding Gibbs measure. In particular, it holds almost surely that $\phi_0 \in H_x^s(\mathbb{T}^2) \setminus L_x^2(\mathbb{T}^2)$ for all $s < 0$. Since the deterministic critical regularity of (NLS) with $(d, p) = (2, 3)$ is given by $s = 0$, the initial value problem with $\phi(0) = \phi_0$ cannot be solved using only deterministic arguments. In order to utilize probabilistic cancellations, Bourgain wrote the solution as

$$\phi = e^{it\Delta} \phi_0 + \psi, \quad (1.8)$$

where the nonlinear remainder ψ solves the forced nonlinear Schrödinger equation

$$\begin{cases} i\partial_t \psi + \Delta \psi =: |e^{it\Delta} \phi_0 + \psi|^2 (e^{it\Delta} \phi_0 + \psi): & (t, x) \in \mathbb{R} \times \mathbb{T}^2, \\ \psi|_{t=0} = 0. \end{cases} \quad (1.9)$$

In the dispersive PDE community, the decomposition (1.8) is known as Bourgain's trick. In the parabolic SPDE community, a similar decomposition is often called Da Prato-Debussche trick [DPD02]. Using a combination of dispersive and probabilistic estimates, Bourgain proved probabilistic nonlinear smoothing for (1.9), which yields control over ψ at the higher³ regularity $s = 1/2-$.

The Gibbs measure problem for the two-dimensional (NLS) with higher-order nonlinearities, i.e., $p \geq 5$, was solved only recently by Deng-Nahmod-Yue [DNY19]. In this case, the deterministic threshold for local well-posedness is $s = 1 - 2/(p - 1)$,

² In dimension $d = 1$, the renormalization in (1.7) is not necessary.

³ In [Bou96], the contraction argument is actually performed at a regularity strictly between 0 and $1/2$, but a minor variant yields the optimal regularity $1/2-$ for ψ (cf. [CLS21, DNY19]).

which is higher than the regularity of the nonlinear remainder in Bourgain's trick. The main contribution of [DNY19] is a more detailed random expansion, which is written as

$$\phi = e^{it\Delta}\phi_0 + \sum_N h^N(t)(e^{it\Delta}P_N\phi_0) + z. \quad (1.10)$$

Here, $(h^N(t))_N$ is a sequence of random averaging operators, $(P_N)_N$ is the sequence of Littlewood-Paley projections, and z is a nonlinear remainder. The term $h^N(t)(e^{it\Delta}P_N\phi_0)$ incorporates certain low $\times \dots \times$ low \times high-interactions between the solution ϕ and the linear evolution of $P_N\phi_0$, which are the worst interactions in (1.7). An essential feature of the argument in [DNY19], which was first observed in a different context by the first author [Bri21], is that the random averaging operator $h^N(t)$ can be chosen as probabilistically independent from the high-frequency data $P_N\phi_0$. In comparison with Bourgain's trick (1.8), the advantage of the random expansion (1.10) is that the nonlinear remainder z lives at regularity 1—, which is above the deterministic threshold for local well-posedness. More recently, the random averaging operators were generalized to random tensors in [DNY20].

In dimension $d = 3$, the Gibbs measure problem for the (NLS) is still open. However, the Gibbs measure problem has been solved for other three-dimensional dispersive equations in [Bri20, BDNY22, DNY21, OOT20, OOT21]. Since a more detailed discussion of this open problem is beyond the scope of this introduction, we refer the interested reader to [DNY21, Section 1.2.1].

1.2.2. The wave maps equation The wave maps equation is a prime example of a semi-linear geometric wave equation. It is the natural generalization of the linear wave equation on $(1 + d)$ -dimensional Minkowski space for scalar-valued fields to fields that map into a Riemannian manifold. Its nonlinearity arises from the geometric constraints imposed by the target manifold and it contains the null form $\partial_\mu\phi^i\partial^\mu\phi^j$. The latter has a favorable algebraic structure, which can be thought of as a cancellation property for the worst interactions of parallel waves.

For sufficiently regular initial data the local well-posedness of the wave maps equation in any space dimension $d \geq 1$ can be established using just energy estimates. At lower regularities the null structure of the wave maps nonlinearity plays a key role for the local existence theory. The optimal sub-critical local well-posedness in $(H_x^s \times H_x^{s-1})(\mathbb{R}^d)$, $s > \frac{d}{2}$, for $d \geq 2$ was obtained by Klainerman-Machedon [KM93, KM95, KM97] and by Klainerman-Selberg [KS97, KS02]. In one space dimension $d = 1$ the analogous optimal sub-critical local well-posedness in $(H_x^s \times H_x^{s-1})(\mathbb{R}^d)$, $s > \frac{1}{2}$, was established by Machihara-Nakanishi-Tsugawa [MNT10]. The initial value problem for the wave maps equation is expected to be ill-posed in $(H_x^s \times H_x^{s-1})(\mathbb{R}^d)$ for $s \leq \frac{d}{2}$, see Tao [Tao00] and references therein.

In one space dimension the global existence of finite energy solutions to the wave maps equation just follows from the local well-posedness in $(H_x^1 \times L_x^2)(\mathbb{R})$ and from energy conservation. In contrast, the global regularity question, i.e., the global existence of smooth solutions, becomes much more subtle in higher space dimensions $d \geq 2$, even for small data. A key difficulty is that the wave maps nonlinearity is no longer perturbative at the critical regularity $(\dot{H}_x^{\frac{d}{2}} \times \dot{H}_x^{\frac{d}{2}-1})(\mathbb{R}^d)$, more precisely this difficulty stems from certain low-high interactions. An influential idea of Tao [Tao01a, Tao01b] to overcome this issue is to exploit the gauge freedom of the wave maps problem, specifically the freedom in the choice of coordinates on the target manifold or rather the freedom in

the choice of frames on the tangent space of the target manifold, to define a physical space gauge transform that recasts the wave maps nonlinearity into a perturbative form. Using this insight, Tao established global regularity for wave maps into the unit sphere for smooth initial data that is small in the critical Sobolev space, first in dimensions $d \geq 5$ in [Tao01a] and then for all dimensions $d \geq 2$ in [Tao01b]. These works also incorporated a functional framework introduced in earlier global regularity results of Tataru [Tat98, Tat01] for smooth initial data that are small in the critical homogeneous Besov space. Other target manifolds were considered in [KR01, NSU03, SS02, Kri03, Kri04, Tat05].

In the energy-critical case of $d = 2$ space dimensions, the dynamics of wave maps with large energies is quite well-understood by now, while in the energy super-critical case $d \geq 3$ little is known about the long-time behavior for large initial data. It turns out that the geometry of the target manifold is decisive for the long-time dynamics of wave maps with large energies in $d = 2$ dimensions. Indeed, the blowup analysis of Struwe [Str03] in the equivariant case uncovered that singularity formation must be tied to the existence of a non-trivial harmonic map into the target manifold. Krieger–Schlag–Tataru [KST08], Rodnianski–Sterbenz [RS10], and Raphaël–Rodnianski [RR12] later constructed examples of (equivariant) wave maps into the unit sphere that blow up in finite time by concentration of a non-trivial harmonic map. Finally, the threshold conjecture for energy-critical wave maps asserts that global regularity is expected for initial data with energy below the energy of any non-trivial finite energy harmonic map into the target manifold. This conjecture was proved independently by Krieger–Schlag [KS12] for the hyperbolic plane as the target, by Tao [Tao08] for all hyperbolic spaces, and by Sterbenz–Tataru [ST10a, ST10b] for all target manifolds that can be isometrically embedded into Euclidean space.

1.2.3. Singular parabolic SPDEs While the subject of this paper is the wave maps equation (WM), which is *hyperbolic*, our methods are influenced by recent advances on singular *parabolic* SPDEs. Since a complete discussion of the literature is beyond the scope of this introduction, we only provide a broad overview and start with the *scalar-valued* setting. While the methods are more general, the reader may think of the parabolic Φ_3^4 -model, which is given by

$$\partial_t \phi - \Delta \phi = - : \phi^3 : + \infty \cdot \phi + \xi \quad (t, x) \in \mathbb{R} \times \mathbb{T}^3. \quad (1.11)$$

Here, the term “ $\infty \cdot \phi$ ” denotes a further renormalization, which goes beyond the Wick-ordering $: \phi^3 :$, and ξ denotes space-time white noise. The local well-posedness of (1.11) was first proven by Hairer in his seminal work on the *theory of regularity structures* [Hai14]. Alternative approaches to singular parabolic SPDEs, such as the parabolic Φ_3^4 -model, are given by the para-controlled calculus of Gubinelli, Imkeller, and Perkowski [GIP15], the renormalization group approach of Kupiainen [Kup16], and an approach of Otto and Weber [OW19]. Out of these four different approaches, the para-controlled calculus of [GIP15] is closest to both the methods in this paper and the dispersive PDE literature. In fact, it served as an inspiration for the random averaging operators in [DNY19], which were previously discussed in Sect. 1.2.1.

We now leave the scalar-valued setting and consider *geometric* equations. In [BGHZ21, Hai16], Bruned, Gabriel, Hairer, and Zambotti studied the geometric stochastic heat equation, which is given by

$$\partial_t \phi_\varepsilon^k = \partial_x^2 \phi_\varepsilon^k + \Gamma_{ij}^k(\phi_\varepsilon) \partial_x \phi_\varepsilon^i \partial_x \phi_\varepsilon^j + h^k(\phi_\varepsilon) + \sigma_i^k(\phi_\varepsilon) \xi_\varepsilon^i + V_{\rho, \sigma}^k(\phi_\varepsilon) \quad (t, x) \in (0, \infty) \times \mathbb{T}. \quad (1.12)$$

Here, (Γ_{ij}^k) are the Christoffel symbols, h is a given vector field, and (σ_i) are vector fields which are chosen depending on the Riemannian metric g . The stochastic forcing ξ_ε^i is a smooth approximation of a space-time white noise ξ^i obtained through convolution with $\rho_\varepsilon(t, x) := \varepsilon^{-3} \rho(t/\varepsilon^2, x/\varepsilon)$. Finally, $V_{\rho, \sigma}^k$ is a vector field which serves as a renormalization and can depend on the vector fields σ_i and the convolution kernel ρ . The main theorem [BGHZ21, Theorem 1.6] proves that the limit $\phi := \lim_{\varepsilon \rightarrow 0} \phi_\varepsilon$ exists on a small time-interval. Furthermore, it proves that there exists a choice $V_{\rho, \sigma} = V_{\rho, \sigma}^{\text{canon}}$ such that the limit ϕ does not depend on the choice of σ or ρ . It is natural to call the corresponding limit the solution of

$$\partial_t \phi^k = \partial_x^2 \phi^k + \Gamma_{ij}^k(\phi) \partial_x \phi^i \partial_x \phi^j + h^k(\phi) + \sigma_i^k(\phi) \xi^i \quad (t, x) \in (0, \infty) \times \mathbb{T}. \quad (1.13)$$

The invariance of Brownian loops under the dynamics of (1.13) is still open and we refer the reader to [BGHZ21, Section 4.3] for a detailed discussion.

In related works, Shen [She21] and Chandra, Chevyrev, Hairer, Shen [CCHS20, CCHS22] obtained similar results for the stochastic Yang mills equation. The main result of this paper (Theorem 1.4), which will be discussed momentarily, is a first step towards extending the results of [BGHZ21, CCHS20, CCHS22, She21] to hyperbolic equations.

1.3. Main result and proof ideas. In most earlier results on wave maps, the initial data is placed in the L^2 -based Sobolev spaces $H_x^s \times H_x^{s-1}$. For our purposes, it is more convenient to work in the L^∞ -based Hölder spaces $C_x^s \times C_x^{s-1}$ (see Definition 2.2). There are two reasons for this: First, the Brownian path B and the white noise velocity V have the same regularity in both Sobolev and Hölder spaces. Second, the $(1+1)$ -dimensional linear wave equation is bounded on Hölder spaces, which is not the case for the $(1+d)$ -dimensional linear wave equation in higher dimensions.

While the main focus of this paper concerns the wave maps equation with random initial data, our estimates also lead to the following theorem on deterministic well-posedness.

Theorem 1.2 (Deterministic well-posedness and mild ill-posedness). *Let (\mathcal{M}, g) be a compact Riemannian manifold and let $r \in \mathbb{R}$.*

- (i) (Well-posedness) *If $r > 1/2$, the wave maps equation (WM) is locally well-posed in $C_x^r \times C_x^{r-1}$.*
- (ii) (Mild ill-posedness) *If $r \leq 1/2$, $\mathcal{M} = \mathbb{S}^{D-1} \subseteq \mathbb{R}^D$, and $D \geq 2$, then the first Picard iterate of the wave maps equation is unbounded on $C_x^r \times C_x^{r-1}$.*

Since the scaling-critical regularity of (WM) in Hölder spaces is given by $r = 0$, the mild ill-posedness for $r \leq 1/2$ does not stem from the scaling symmetry. Instead, it is a result of bad high \times high \rightarrow low-interactions in the nonlinearity. Due to the mild ill-posedness for $r \leq 1/2$, we cannot treat the wave maps equation with our random initial data (B, V) using a deterministic contraction-mapping argument.

Remark 1.3. Theorem 1.2 only yields a rather mild form of ill-posedness. It would therefore be interesting to also prove stronger forms of ill-posedness such as the failure of continuous dependence on the initial data or even norm inflation. Norm inflation may however be difficult to establish in $C_x^r \times C_x^{r-1}$ for $0 < r \leq 1/2$ since the geometric constraints $\phi(t, x) \in \mathcal{M}$ prevent arbitrary growth in the L_x^∞ -norm. Since the main focus of this article is on geometric and probabilistic aspects of the well-posedness theory, we leave the proof of stronger forms of ill-posedness as an open problem (and refer the reader to [Kis19, For20, FO20, Oh17, ST20] for more detailed discussions of ill-posedness).

Before we state the rigorous version of our main result (Theorem 1.4), we give a precise definition of the random data. To this end, we first let $W: \mathbb{R} \rightarrow \mathbb{R}^D$ be a Euclidean Brownian motion in the ambient space and we fix a reference point $B_0 \in \mathcal{M}$. For each $p \in \mathcal{M}$, we let $P(p): \mathbb{R}^D \rightarrow T_p \mathcal{M}$ be the orthogonal projection onto the tangent space of \mathcal{M} at p . Then, the Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ is defined as the solution to the Stratonovich SDE

$$dB(x) = P(B(x)) \circ dW(x), \quad B(0) = B_0 \in \mathcal{M}. \quad (1.14)$$

In addition to the Brownian path B itself, we also define smooth approximations $(B^\varepsilon)_\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$. We implicitly restrict the parameter ε to dyadic numbers, but do not further reflect this in our notation. We then first define the smooth approximations $(W^\varepsilon)_\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^D$ of the Euclidean Brownian motion W by

$$W^\varepsilon := P_{\leq \varepsilon^{-1}} W,$$

where $P_{\leq \varepsilon^{-1}}$ is the Littlewood-Paley projection from Definition 2.1 below. Then, we define the smooth path $B^\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$ as the solution to the classical ODE

$$\partial_x B^\varepsilon(x) = P(B^\varepsilon(x)) \partial_x W^\varepsilon(x), \quad B^\varepsilon(0) = B_0 \in \mathcal{M}. \quad (1.15)$$

In Corollary 2.34 below, it is shown that the smooth paths $(B^\varepsilon)_\varepsilon$ converge to the Brownian path B in C_{loc}^s for $s < 1/2$. It remains to define the white noise velocity V and its smooth approximations $(V^\varepsilon)_\varepsilon$. To this end, we let $\bar{W}: \mathbb{R} \rightarrow \mathbb{R}^D$ be an independent copy of W and define $\bar{W}^\varepsilon := P_{\leq \varepsilon^{-1}} \bar{W}$. Then, we explicitly define

$$V^\varepsilon(x) := P(B^\varepsilon(x)) \partial_x \bar{W}^\varepsilon(x). \quad (1.16)$$

Due to the projection $P(B^\varepsilon(x))$, it holds that $V^\varepsilon \in (B^\varepsilon)^* T \mathcal{M}$, i.e., $V^\varepsilon(x) \in T_{B^\varepsilon(x)} \mathcal{M}$ for all $x \in \mathbb{R}$. In Corollary 2.34 below, it is shown that $(V^\varepsilon)_\varepsilon$ converges in C_{loc}^{s-1} for $s < 1/2$. Therefore, we can define the white noise velocity V as

$$V := \lim_{\varepsilon \rightarrow 0} V^\varepsilon. \quad (1.17)$$

Equipped with the Brownian path B , the white noise velocity V , and their smooth approximations, we can now state our main result.

Theorem 1.4 (Probabilistic local well-posedness). *Let $B: \mathbb{R} \rightarrow \mathcal{M}$ be the Brownian path, let $V \in B^* T \mathcal{M}$ be the white noise velocity, and let $(B^\varepsilon)_{\varepsilon > 0}$ and $(V^\varepsilon)_{\varepsilon > 0}$ be their smooth approximations. Then, for all $\tau > 0$ and $R \geq 1$, there exists an event $\mathcal{E}(\tau, R)$ such that the following two properties hold:*

(i) (“High”-probability) *We have that*

$$\mathbb{P}(\mathcal{E}(\tau, R)) \geq 1 - CR \exp(-c\tau^{-c}),$$

where $C = C(\mathcal{M}) \geq 1$ and $c > 0$ are constants.

(ii) (LWP) *On the event $\mathcal{E}(\tau, R)$, the smooth global solutions ϕ^ε of (WM) with initial data $\phi^\varepsilon[0] = (B^\varepsilon, V^\varepsilon)$ converge in $(C_t^0 C_x^s \cap C_t^1 C_x^{s-1})([-\tau, \tau] \times [-R, R] \rightarrow \mathcal{M})$.*

We note that the local well-posedness statement in Theorem 1.4 is slightly non-standard since, in order for $\mathbb{P}(\mathcal{E}(\tau, R))$ to be close to one, τ needs to be small depending on R . Put differently, Theorem 1.4 does not yield local well-posedness in time, but rather yields local well-posedness in space-time. Since the random data (B, V) does not decay in space and the wave maps equation exhibits finite speed of propagation, this version of local well-posedness is natural.

We now describe the main ideas in the proof of Theorem 1.4. Throughout this informal discussion, we denote the regularity parameter for the Brownian path by $s < 1/2$. Furthermore, we formally set $\varepsilon = 0$ and omit smooth cut-off functions from our notation. The first step, which was already used in [KT98, MNT10, Tao00], is to switch from Cartesian to null-coordinates. We define

$$u := x - t \quad \text{and} \quad v := x + t.$$

Due to d'Alembert's formula, the linear evolution of the initial data (B, V) is given by⁴

$$\phi_{\text{lin}}(u, v) = \phi^+(u) + \phi^-(v), \quad (1.18)$$

where

$$\phi^\pm(x) = \frac{1}{2} \left(B(x) \mp \int_0^x dy V(y) \right). \quad (1.19)$$

We refer to ϕ^+ and ϕ^- as the right and left-moving linear waves, respectively. The wave maps equation in null coordinates is given by

$$\begin{cases} \partial_u \partial_v \phi^k = -\Pi_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \\ \phi|_{u=v} = B, \quad (\partial_v - \partial_u)\phi|_{u=v} = V. \end{cases} \quad (1.20)$$

The most difficult aspects of our argument are linked to the absence of nonlinear smoothing for (1.20). In particular, we cannot rely on Bourgain's trick. Instead, we require a more delicate Ansatz, which is related to but different from the random averaging operators in (1.10). To motivate this Ansatz, we heuristically discuss low \times high and high \times high \rightarrow low-interactions.

1.3.1. Low \times high-interactions Our treatment of the low \times high-interactions is motivated by the gauge transform of Tao [Tao01a, Tao01b], which was already mentioned in Sect. 1.2. While we will also be multiplying the (linear) evolution with a low-frequency modulation, our construction of the modulation differs significantly from [Tao01a, Tao01b].

We first recall from (1.18) that the linear evolution is given by $\phi_{\text{lin}}^i(u, v) = \phi^{+,i}(u) + \phi^{-,i}(v)$. We now focus on the right-moving component $\phi^{+,i}(u)$ and restrict to frequencies of size $\sim M$. The corresponding portion of the linear evolution is given by $\phi_M^{+,i}(u) := (P_M \phi^{+,i})(u)$. More generally, let us replace $\phi_M^{+,i}(u)$ by the modulated right-moving wave $A_{M,m}^{+,i}(u, v) \phi_M^{+,m}(u)$. For the moment, we assume that the modulation

⁴ For technical reasons, we will later write the linear evolution as $\phi_{\text{lin}} = \theta(\phi^+(u) + \phi^-(v))$, where $\theta > 0$ is a small parameter. Due to this, we will need to adjust the definitions of ϕ^+ and ϕ^- , see e.g. (3.2). In the introduction, we ignore this technicality.

$A_{M,m}^{+,i}$ is supported on frequencies of size ~ 1 in both variables. Inserting this component into the nonlinearity yields

$$-\mathbb{I}_{ij}^k(\phi(u, v))\partial_u(A_{M,m}^{+,i}(u, v)\phi_M^{+,m}(u))\partial_v\phi^j(u, v).$$

Since we are currently discussing low \times high-interactions, we turn to the sub-term

$$-P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j)(u, v)\partial_u(A_{M,m}^{+,i}(u, v)\phi_M^{+,m}(u)), \quad (1.21)$$

where the product $\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j$ enters at low frequencies in both variables. Eventually, we will solve the wave maps equation (1.20) using its Duhamel integral formulation. Using explicit calculations (see Proposition 2.23), the Duhamel integral of (1.21) is of the form

$$\begin{aligned} & \text{Duh} \left[-P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j)(u, v)\partial_u(A_{M,m}^{+,i}(u, v)\phi_M^{+,m}(u)) \right] \\ &= -\left(\int_u^v dv' P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j)(u, v')A_{M,m}^{+,i}(u, v') \right) \phi_M^{+,m}(u) + \{\text{error terms}\}. \end{aligned} \quad (1.22)$$

The leading term in (1.22) has the same regularity as $\phi_M^+(u)$, which witnesses the absence of nonlinear smoothing. In particular, (1.22) cannot be absorbed into a smoother nonlinear remainder. However, the leading term in (1.22) is exactly of the same form as our starting point $A_{M,m}^{+,i}(u, v)\phi_M^{+,m}(u)$. Therefore, we can hope to absorb it into the k -th component $A_{M,m}^{+,k}(u, v)\phi_M^{+,m}(u)$. By pursuing this idea, one is quickly lead to the ordinary differential equation (ODE) given by

$$\partial_v A_{M,m}^{+,k}(u, v) = -P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j)(u, v)A_{M,m}^{+,i}(u, v), \quad A_{M,m}^{+,k}(u, u) = \delta_m^k. \quad (1.23)$$

The initial value $A_{M,m}^{+,k}(u, u) = \delta_m^k$ in (1.23) is due to the linear evolution, which has to be contained in $A_{M,m}^{+,k}(u, v)\phi_M^{+,m}(u)$. Provided that the pre-factor $\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j$ is well-defined, the ODE (1.23) can easily be solved using the Picard-Lindelöf theorem. Unfortunately, (1.21) is an oversimplification of all low \times high-interactions. As a result, (1.23) does not present itself as a suitable model for the modulation equations (see Sect. 8). This is due to a further culprit hidden in $\mathbb{I}_{ij}^k(\phi)\partial_v\phi^j$, which is given by

$$P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi))(u, v)\partial_v\phi_N^{-,j}(v). \quad (1.24)$$

Here, ϕ^- is the left-moving component of the initial data and we take $1 \leq N \leq M$.⁵ Restricting our attention only to the contribution of (1.24), we are led to the ODE

$$\partial_v A_{M,m}^{+,k}(u, v) = -P_{\lesssim 1}^{u,v}(\mathbb{I}_{ij}^k(\phi))(u, v)A_{M,m}^{+,i}(u, v)\partial_v\phi_N^{-,j}(v). \quad (1.25)$$

It is evident from (1.25) that we should no longer view the modulation $A_{M,m}^{+,k}$ as being supported on frequencies of size ~ 1 in the v -variable. Unfortunately, (1.25) cannot be solved using classical theory for ODEs. The reason is that ϕ^- only has regularity s , which suggests that $v \mapsto A_M^+(u, v)$ also only has regularity s . Since $s < 1/2$, the regularity information is insufficient to even define the right-hand side in (1.25). Instead of classical ODE methods, we utilize the para-controlled approach to rough ODEs by Gubinelli,

⁵ When $N > M$, the roles of ϕ^+ and ϕ^- should be reversed.

Imkeller, and Perkowski [GIP15]. Due to certain high \times high \rightarrow low-interactions, which will be described in Sect. 1.3.2, it is convenient to separate the cases $1 \leq N \leq M^{1-\delta}$ and $M^{1-\delta} < N \leq M$. After reversing the roles of the u and v -variables, we are lead to the following four terms in our Ansatz:

- (1) The modulated right-moving wave $A_{M,m}^{+,i}(u, v)\phi_M^{+,m}(u)$.
- (2) The modulated left-moving wave $A_{N,n}^{-,i}(u, v)\phi_N^{-,n}(v)$.
- (3) The bilinear term $B_{M,N,mn}^i(u, v)\phi_M^{+,m}(u)\phi_N^{-,n}(v)$, which contains left and right-moving components and only occurs when $\min(M, N) \geq \max(M, N)^{1-\delta}$.
- (4) The nonlinear remainder $\psi^i(u, v)$, which lives at a higher regularity.

1.3.2. High \times high \rightarrow low-interactions As stated in Theorem 1.2, the one-dimensional wave maps equation is (deterministically) ill-posed in C^r for all $r \leq 1/2$, which is a result of high \times high \rightarrow low-interactions. This is in sharp contrast to the deterministic theory for wave maps in high dimensions [NSU03, SS02, Tao01a], where high \times high \rightarrow low-interactions are relatively harmless. In order to go beyond the deterministic theory, we need to rely on probabilistic cancellations for the Brownian path B and the white noise velocity V . The main ingredient is the high \times high \rightarrow low-estimate

$$\mathbb{E} \left[\max_{\pm 1, \pm 2} \max_{1 \leq i, j \leq D} \sup_{M \sim N} M^s \left\| P_M \phi^{\pm 1, i}(x) \partial_x P_N \phi^{\pm 2, j}(x) \right\|_{C^{s-1}} \right] \lesssim 1, \quad (1.26)$$

where $\pm 1, \pm 2 \in \{+, -\}$. We emphasize that since $s + (s - 1) < 0$, the left-hand side in (1.26) cannot be bounded using only that $\phi^\pm \in C^s$. In our analysis, we then encounter two different forms of frequency-resonances.

Frequency-resonances involving only the linear waves ϕ^\pm : These terms cannot even be defined using only deterministic estimates, but are relatively harmless once the probabilistic ingredient (1.26) is taken into account. To illustrate this, let $K \sim M \sim N$ be comparable frequency scales and consider the cubic term

$$P_{\lesssim 1}^u(\phi_K^{+,k}(u) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v). \quad (1.27)$$

The term (1.27) naturally occurs after a para-linearization of the second fundamental form $\mathbb{I}(\phi)$ in (1.20). It has u -frequency ~ 1 , v -frequency $\sim N$, and, due to (1.26), the amplitude

$$\left\| P_{\lesssim 1}^u(\phi_K^{+,k}(u) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v) \right\|_{L_{u,v}^\infty} \lesssim M^{-s} N^{1-s} \sim N^{1-2s}. \quad (1.28)$$

As our analysis will show (see Lemma 4.4 and Proposition 5.3), (1.28) is sufficient to treat the cubic term (1.27) as a smooth remainder.

Frequency-resonances involving the linear ϕ^\pm and the remainder ψ : These terms can be defined using only deterministic ingredients. Despite being well-defined, however, the resulting estimates of the resonant term are worse than (1.28).

Let $r \in (1/2, 1)$ denote the regularity of the smooth remainder $\psi \in C_u^r C_v^r$ (see Definition 2.2 below). Similar as in (1.27), we let $K \sim M \sim N$ be frequency scales and consider the resonant term

$$P_{\lesssim 1}^u(P_K^u \psi^k(u, v) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v). \quad (1.29)$$

In comparison with (1.27), we only replaced $\phi_K^{+,k}(u)$ by $P_K^u \psi^k(u, v)$. We make no restrictions on the v -frequencies of $\psi^k(u, v)$, and the worst case corresponds to v -frequencies of size ~ 1 . Since $\psi \in C_u^r C_v^r$ is arbitrary, we can only use direct (deterministic) estimates of (1.29), which yield

$$\|P_K^u (P_K^u \psi^k(u, v) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v)\|_{L_{u,v}^\infty} \lesssim K^{-r} M^{1-s} N^{1-s} \sim N^{1-2s+1-r}. \quad (1.30)$$

Since $r < 1$, the estimate is worse than (1.28). More importantly, it only holds that

$$\begin{aligned} & \|P_{\lesssim 1}^u (P_K^u \psi^k(u, v) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v)\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim 1^{r-1} N^{r-1} \|P_{\lesssim 1}^u (P_K^u \psi^k(u, v) \partial_u \phi_M^{+,m}(u)) \partial_v \phi_N^{-,n}(v)\|_{L_{u,v}^\infty} \\ & \lesssim N^{1-2s}. \end{aligned} \quad (1.31)$$

Since $s < 1/2$, the right-hand side of (1.31) is unbounded in N . As a result, (1.29) cannot be absorbed back into the nonlinear remainder ψ . Instead, (1.29) forms an additional contribution to the modulated left-moving wave $A_{N,n}^{-,i}(u, v) \phi_N^{-,n}(v)$. Unfortunately, the high \times high \rightarrow low-interactions between ϕ^\pm and ψ do not only occur in (1.29), but are part of many different terms in $\Pi_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j$.

We remark that (1.30) can be improved if the condition $N \sim M$ is replaced by $N \leq M^{1-\delta}$. This is the reason for introducing the bilinear term $B_{M,N,mn}^i(u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v)$, which isolates the problematic case $N > M^{1-\delta}$.

1.3.3. Further remarks We now make further remarks and a few comparisons of this article with related works.

- (i) While the $(1+1)$ -dimensional wave maps equation is completely integrable (see e.g. [KT98, Poh76, TU04]), our argument does not rely on complete integrability. In fact, our well-posedness theory also applies to the more general system

$$\partial_\mu \partial^\mu \phi^k = -\Pi_{ij}^k(\phi) \partial_\mu \phi^i \partial^\mu \phi^j + F^k(\phi),$$

where F is (the extension to \mathbb{R}^D of) a vector field on \mathcal{M} .

- (ii) The approximations $(B^\epsilon)_{\epsilon>0}$ and $(V^\epsilon)_{\epsilon>0}$ from (1.15) and (1.16) and the corresponding wave maps $(\phi^\epsilon)_{\epsilon>0}$ depend on the convolution kernel ρ , which is used in the definition of our Littlewood-Paley operators (see Definition 2.1). However, we expect that their limits B , V , and ϕ do not depend on the precise choice of ρ . That is, we expect that any convolution kernel $\bar{\rho}$ satisfying the conditions in Definition 2.1 leads to the same limit. This should follow from variants of Lemma 2.30.(ii), Proposition 2.32.(vi), and Proposition 2.33.(vi), whose proof should allow us to control the differences between stochastic objects based on ρ and $\bar{\rho}$.
- (iii) In this article, we view the compact Riemannian manifold \mathcal{M} as an isometrically embedded submanifold of a Euclidean space \mathbb{R}^D . But neither the isometric embedding nor the Euclidean space \mathbb{R}^D are unique, and it is an interesting problem to show that the (law of the) limit ϕ does not depend on them.
- (iv) The modulated linear wave $A_{N,n}^{-,k}(u, v) \phi_N^{-,n}(v)$ shares similarities with the adapted linear evolutions in [Bri21] and the random averaging operators in [DNY19], see Sect. 1.2.1. However, the modulation A_N^- and linear wave ϕ_N^- are probabilistically dependent. There are two reasons for this: First, the geometric constraints on the

Brownian path B and white noise velocity V create probabilistic dependencies between low and high frequencies. Second, the high \times high \rightarrow low-interaction in (1.30), which enters into the modulation A_N^- , depends on high frequency terms in the initial data. To compensate for the lack of probabilistic independence, however, the linear map $\phi_N^- \mapsto A_{N,n}^{-,k}(u, v)\phi_N^{-,n}(u)$ consists of a simple multiplication, which is easier to handle than the random averaging operators in (1.10).

- (v) In [KLS20], two of the authors in joint work with J. Krieger obtained a probabilistic small data global existence result for the energy-critical Maxwell-Klein-Gordon equation relative to the Coulomb gauge. It is the first global existence result for a geometric wave equation for random initial data at super-critical regularity. Similarly to [Bri21], the proof is based upon an induction on frequency procedure and an adapted linear-nonlinear decomposition. The latter relies on a delicate global parametrix construction. It is worth noting that already in the deterministic study of the global well-posedness of the energy-critical Maxwell-Klein-Gordon equation, certain low-high interactions are non-perturbative at critical regularity and have to be incorporated into the linear operator. In comparison to [KLS20], a novel feature of this work is that the random data arises naturally from geometric considerations.
- (vi) In [BR20], Brzeźniak and Rana consider a stochastic wave maps equation. In null-coordinates on Minkowski space \mathbb{R}^{1+1} and intrinsic coordinates on the Riemannian manifold \mathcal{M} , it is given by

$$\partial_u \partial_v \phi^k = -\Gamma_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j + \frac{1}{4} \sigma_i^k(\phi) \partial_u \partial_v \Xi^i. \quad (1.32)$$

Here, $\Gamma_{ij}^k : M \rightarrow \mathbb{R}$ are the Christoffel symbols, $\sigma_i^k : \mathcal{M} \rightarrow \mathbb{R}$ are smooth functions, and the stochastic terms $\Xi^i : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}$ are given by fractional Brownian sheets with Hurst indices $3/4 < H_1, H_2 < 1$. The main result [BR20, Theorem 4.3], which builds on [MNT10], yields the local well-posedness of the stochastic wave maps equation (1.32). It would be interesting to see if the methods in this paper could be used to extend their result to Hurst indices $s < H_1, H_2 < 1$, where $s = 1/2-$.

- (vii) In [BJ22], which appeared after the preprint of this manuscript, Brzeźniak and Jendrej studied lattice approximations of wave maps into spheres with Brownian initial data. The main result of [BJ22] shows that, as the lattice spacing tends to zero, a subsequence of the discretized solutions converges in law and that the law of the limit is invariant under time-like translations. The main differences between [BJ22, Theorem 1] and Theorem 1.4 are that [BJ22, Theorem 1] is global (rather than local) in time, but Theorem 1.4 concerns strong (rather than weak) solutions.
- (viii) As briefly mentioned in Sect. 1.2.3, Theorem 1.4 is a first step towards extending the results in [BGHZ21] from the parabolic to the hyperbolic setting. The most intriguing difference between [BGHZ21] and our work is the use of renormalization techniques, which are not needed in the proof of Theorem 1.4. Such techniques, however, may be very relevant in the open problem described below, which requires a more detailed understanding of finite-dimensional approximations of (WM).

1.4. Open problem: invariant Gibbs measure. Our original motivation to study the $(1+1)$ -dimensional wave maps equation was the following problem, which remains unsolved.

Geometric Gibbs measure problem. *Prove the existence and invariance of the Gibbs measure for the wave maps equation (WM).*

As discussed in Sect. 1.1, the Gibbs measure for the wave maps equation has not yet been constructed. Aside from the construction of the measure, we encountered another (substantial) difficulty in our attempt to solve this problem. Previous proofs of invariance (see e.g. [Bou94, Bou96]) all rely on finite-dimensional approximations of the full system. In the derivation of the finite-dimensional approximation, however, one has to carefully preserve the important properties of the full system. In the case of the wave maps equation, the finite-dimensional approximation should still exhibit the following two structures:

- (i) The null structure.
- (ii) The Hamiltonian structure.

It is easy to derive finite-dimensional approximations which preserve either of these two structures, but this seems to be insufficient to solve the geometric Gibbs measure problem. So far, we were unable to derive a finite-dimensional truncation which preserves both structures simultaneously. For related discussions of finite-dimensional approximations of random dispersive equations and singular stochastic PDEs, we refer the reader to [DTV15, NORBS12] and [CM18, EH19, FH17, HM12, HM18], respectively.

2. Preparations

2.1. Notation and parameters. Let $d \geq 1$ and let $f \in \mathcal{S}(\mathbb{R}^d)$ be a Schwartz function. We define the Fourier and inverse Fourier transform of f by

$$\widehat{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} dx e^{-ix\xi} f(x) \quad \text{and} \quad \check{f}(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} d\xi e^{+ix\xi} f(\xi).$$

In the following, we often use dyadic decompositions of frequency space and we refer to the corresponding dyadic scales simply as frequency scales. The frequency scales will be denoted exclusively by capital letters such as K , L , M , and N . For two frequency scales M and N , we define

$$\begin{aligned} M \ll N &: \Longleftrightarrow M \leq 2^{-10} N, \\ M \lesssim N &: \Longleftrightarrow M \leq 2^{10} N, \\ M \sim N &: \Longleftrightarrow 2^{-10} N < M < 2^{10} N. \end{aligned} \tag{2.1}$$

With a slight abuse of notation, we also use “ \ll ” and “ \lesssim ” for quantities other than frequency scales, but where the exact definition differs slightly from (2.1). Let C and c be sufficiently large and small absolute constants, respectively. For any $A, B > 0$, we write $A \ll B$ if $A \leq cB$ and $A \lesssim B$ if $A \leq CB$. In the following, the precise meaning of “ \ll ” and “ \lesssim ” will always be clear from the context.

We now turn to the parameters used in our analysis. The main parameters are given by $r, s, \delta \in \mathbb{R}$, which satisfy

$$0 < \frac{1}{2} - s \ll \delta \ll \frac{3}{4} - r \ll 1. \tag{2.2}$$

The parameter s represents the (low) regularity of the initial data and the parameter r represents the (high) regularity of the smoother remainder. The parameter δ is used in our definition of a modified low-high para-product. For notational convenience, we also choose

$$\sigma := 100\delta. \quad (2.3)$$

In addition, we choose a (less significant) parameter η satisfying

$$0 < \frac{1}{2} - s \ll \eta \ll \delta. \quad (2.4)$$

The parameter η will mostly be used to ensure the summability over dyadic scales. For notational convenience, we also define

$$r' = r - \eta. \quad (2.5)$$

During the first reading of the paper, we encourage the reader to mentally replace $s \rightarrow 1/2$, $\eta \rightarrow 0$, $\delta \rightarrow 0$, and $r', r \rightarrow 3/4$.

Finally, we let $\theta > 0$ be a small parameter which is allowed to depend on s , η , δ , and r . It will be used as an upper bound on the size of the re-scaled initial data and therefore plays a different role than the other parameters, which are linked to frequency-scale restrictions and/or regularities.

Finally, we define modifications of (2.1), which involve the parameter $\delta > 0$. For all frequency scales M and N , we define

$$\begin{aligned} M \ll_{\delta} N &: \Longleftrightarrow M \leq N^{1-\delta}, \\ M \lesssim_{\delta} N &: \Longleftrightarrow M \leq N^{1/(1-\delta)}, \\ M \sim_{\delta} N &: \Longleftrightarrow N^{1-\delta} < M < N^{1/(1-\delta)}. \end{aligned} \quad (2.6)$$

In the extrinsic formulation of the wave maps equation, we previously encountered the second fundamental form $\mathbb{I}: T\mathcal{M} \times T\mathcal{M} \rightarrow N\mathcal{M}$. Our assumption that \mathcal{M} is a smooth, compact Riemannian manifold without boundary guarantees the uniform boundedness of the second fundamental form \mathbb{I} and of all of its derivatives in the embedding $\mathcal{M} \hookrightarrow \mathbb{R}^D$. In order to work in the ambient space \mathbb{R}^D , we require an extension $\tilde{\mathbb{I}}$ of \mathbb{I} . The extension $\tilde{\mathbb{I}}$ is determined by its components $\tilde{\mathbb{I}}_{ij}^k: \mathbb{R}^D \rightarrow \mathbb{R}$, which can be chosen to satisfy the following properties:

- (1) For all $\phi \in \mathcal{M}$ and $V, W \in T_{\phi}\mathcal{M}$, it holds that $\mathbb{I}^k(\phi)(V, W) = \tilde{\mathbb{I}}_{ij}^k(\phi)V^iW^j$.
- (2) For all $1 \leq i, j, k \leq D$, it holds that $\tilde{\mathbb{I}}_{ij}^k \in C_c^{\infty}(\mathbb{R}^D \rightarrow \mathbb{R})$ and $\tilde{\mathbb{I}}_{ij}^k = \tilde{\mathbb{I}}_{ji}^k$.

Such an extension $\tilde{\mathbb{I}}$ of the second fundamental form can be realized as the Hessian of a smooth and compactly supported extension of the nearest point projection map, which under our assumptions is well-defined in a tubular neighborhood of \mathcal{M} , see for instance [Sim96, Section 2.12.3]. With a slight abuse of notation, we now identify \mathbb{I} with $\tilde{\mathbb{I}}$ and simply write \mathbb{I} for both the original second fundamental form and its extension.

Throughout this article, we fix a bump function $\chi \in C_c^{\infty}(\mathbb{R} \rightarrow [0, 1])$ which satisfies

$$\chi(x) = 1 \text{ for all } x \in [-2, 2] \text{ and } \chi(x) = 0 \text{ for all } x \notin [-21/10, 21/10]. \quad (2.7)$$

Furthermore, we define two functions $\chi^+, \chi^-: \mathbb{R}_{u,v}^{1+1} \rightarrow [0, 1]$ by

$$\chi^+(u, v) := \chi(u) \quad \text{and} \quad \chi^-(u, v) := \chi(v). \quad (2.8)$$

2.2. Function spaces, para-products and product estimates. In this section, we define the Bourgain-type space $C_u^{\gamma_1} C_v^{\gamma_2}$. Despite its simplicity, the space $C_u^{\gamma_1} C_v^{\gamma_2}$ is essential for making use of the null-structure in the wave maps equation. After this definition, we turn to para-products and their estimates in the $C_u^{\gamma_1} C_v^{\gamma_2}$ -spaces.

We let $C_b^\infty(\mathbb{R})$ be the space of bounded smooth functions $f: \mathbb{R} \rightarrow \mathbb{R}$ with bounded derivatives, i.e., satisfying

$$\|\partial_x^k f(x)\|_{L^\infty(\mathbb{R})} \leq C(k, f) < \infty$$

for all $k \geq 0$. Furthermore, we let $C_c^\infty(\mathbb{R})$ be the subspace of $C_b^\infty(\mathbb{R})$ consisting of compactly supported smooth functions. We now recall the definition of Littlewood-Paley operators.

Definition 2.1 (*Littlewood-Paley operators*). Let $\rho: \mathbb{R} \rightarrow \mathbb{R}$ be a smooth even cut-off function satisfying $\rho(\xi) = 1$ for all $\xi \in [-7/8, 7/8]$ and $\rho(\xi) = 0$ for all $\xi \notin [-9/8, 9/8]$. We then define

$$\rho_1(\xi) := \rho(\xi) \quad \text{and} \quad \rho_N(\xi) := \rho(\xi/N) - \rho(2\xi/N) \quad \text{for all } N \geq 2. \quad (2.9)$$

For any function $f \in C_b^\infty(\mathbb{R})$, the Littlewood-Paley projections $\{P_N f\}_{N \geq 1}$ are defined as

$$P_N f(x) := (\check{\rho}_N * f)(x), \quad (2.10)$$

where $\check{\rho}_N$ is the inverse Fourier transform of ρ_N . In addition, we define the fattened Littlewood-Paley operators \tilde{P}_N by

$$\tilde{P}_N := \sum_{M: M \sim N} P_M. \quad (2.11)$$

Furthermore, if $f \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$, the Littlewood-Paley projections in the u and v -variables are defined as

$$P_N^u f(u, v) = (\check{\rho}_N *_{u} f)(u, v) \quad \text{and} \quad P_N^v f(u, v) = (\check{\rho}_N *_{v} f)(u, v), \quad (2.12)$$

where $*_u$ and $*_v$ denote the convolution in the u and v -variable, respectively. The fattened Littlewood-Paley projections \tilde{P}_N^u and \tilde{P}_N^v are defined similarly as in (2.11).

Equipped with Definition 2.1, we can now define the Bourgain-type spaces $C_u^{\gamma_1} C_v^{\gamma_2}$.

Definition 2.2 (*Hölder and Bourgain-type spaces*). For any regularity parameter $\gamma \in \mathbb{R}$ and any function $f: \mathbb{R} \rightarrow \mathbb{R}$, we define the C^γ -norm of f by

$$\|f\|_{C^\gamma} = \|f\|_{C^\gamma(\mathbb{R})} := \sup_{N \geq 1} N^\gamma \|P_N f(x)\|_{L_x^\infty(\mathbb{R})}. \quad (2.13)$$

The corresponding Hölder space C^γ is defined as the completion of $C_b^\infty(\mathbb{R})$.

For any two regularity parameters $\gamma_1, \gamma_2 \in \mathbb{R}$ and any function $f: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}$, we define the $C_u^{\gamma_1} C_v^{\gamma_2}$ -norm of f by

$$\|f\|_{C_u^{\gamma_1} C_v^{\gamma_2}} = \|f\|_{C_u^{\gamma_1} C_v^{\gamma_2}(\mathbb{R}_{u,v}^{1+1})} := \sup_{N_1, N_2 \geq 1} N_1^{\gamma_1} N_2^{\gamma_2} \|P_{N_1}^u P_{N_2}^v f(u, v)\|_{L_{u,v}^\infty(\mathbb{R}_{u,v}^{1+1})}. \quad (2.14)$$

Similar as above, we define the corresponding Bourgain-type space $C_u^{\gamma_1} C_v^{\gamma_2}$ as the completion of $C_b^\infty(\mathbb{R}_{u,v}^{1+1})$ under the $C_u^{\gamma_1} C_v^{\gamma_2}$ -norm.

Remark 2.3. We refer to $C_u^{\gamma_1} C_v^{\gamma_2}$ as a Bourgain-type space since it is the natural L^∞ -based analogue of the L^2 -based Bourgain space $X^{s,b}$ (see e.g. [Bou93, Tao06]). Even if $\gamma_1 = \gamma_2 = \gamma \in \mathbb{R}$, the Bourgain-type space $C_u^{\gamma_1} C_v^{\gamma_2}$ does not coincide with the usual Hölder spaces $C^\gamma(\mathbb{R}_{u,v}^{1+1})$. While (2.14) contains the pre-factor $N_1^{\gamma_1} N_2^{\gamma_2}$, the usual Hölder norm would contain $\max(N_1, N_2)^\gamma$. This difference is essential for making use of the null structure in the wave maps equation.

We now turn to the para-product operators. First, we recall the definition of standard low \times high, high \times high, and high \times low para-products (in a single variable). For our purposes, it is convenient to give ourselves more room in the frequency-scales, which leads to the modified para-products below. Finally, we extend the para-products in a single variable to para-products acting on either the u or v -coordinate.

Definition 2.4 (*Para-product operators*). In this definition, we define three different kinds of para-product operators.

(1) Standard para-products: For any $f, g \in C_b^\infty(\mathbb{R})$, we define the para-products

$$f \otimes g := \sum_{M \ll N} P_M f \cdot P_N g, \quad (2.15)$$

$$f \oslash g := \sum_{M \sim N} P_M f \cdot P_N g, \quad (2.16)$$

$$f \oplus g := \sum_{M \gg N} P_M f \cdot P_N g. \quad (2.17)$$

In other words, \otimes , \oslash , and \oplus correspond to the low \times high, high \times high, and high \times low para-product, respectively. Similarly, we define

$$f \lesssim g := \sum_{M \lesssim N} P_M f \cdot P_N g, \quad (2.18)$$

$$f \gtrsim g := \sum_{M \gtrsim N} P_M f \cdot P_N g, \quad (2.19)$$

$$f \not\sim g := \sum_{M \not\sim N} P_M f \cdot P_N g. \quad (2.20)$$

In other words, $f \lesssim g$ contains low \times high and high \times high-interactions, $f \gtrsim g$ contains high \times low and high \times high-interactions, and $f \not\sim g$ contains low \times high and high \times low-interactions.

(2) Modified para-products: Let $\sigma \in (0, 1)$ be as in (2.2) and (2.3). For any $f, g \in C_b^\infty(\mathbb{R})$, we also define modified para-products by

$$f \otimes_\sigma g := \sum_{M \leq N^{1-\sigma}} P_M f \cdot P_N g, \quad (2.21)$$

$$f \oslash_\sigma g := \sum_{M > N^{1-\sigma}} P_M f \cdot P_N g, \quad (2.22)$$

$$f \downarrow_\sigma g := \sum_{M \sim N} \sum_{K \leq \min(M, N)^\sigma} P_K (P_M f \cdot P_N g). \quad (2.23)$$

As a result, \llcorner_σ contains less frequency-interactions than \lrcorner and \gtrcorner_σ contains more frequency-interactions than \gtrcorner . The operator \downarrow_σ contains only high \times high \rightarrow low-interactions and is therefore a modification of \lrcorner .

- (3) Para-products in u and v : Finally, if $f, g \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$, we define para-product operators in individual variables by adding superscripts to our previous para-product operators. For example, the analogues of (2.15), (2.16), and (2.17) are given by

$$(f \llcorner^u g)(u, v) := \sum_{M \ll N} P_M^u f(u, v) \cdot P_N^u g(u, v), \quad (2.24)$$

$$(f \lrcorner^u g)(u, v) := \sum_{M \sim N} P_M^u f(u, v) \cdot P_N^u g(u, v), \quad (2.25)$$

$$(f \gtrcorner^u g)(u, v) := \sum_{M \gg N} P_M^u f(u, v) \cdot P_N^u g(u, v). \quad (2.26)$$

All other para-product operators used in this paper, such as \lrcorner^u , \llcorner_σ^u , and \downarrow_σ^u , are defined similarly.

We now present the basic bilinear estimate in our product spaces. The results and proofs are similar to the standard para-product estimates in Besov or Hölder spaces, see e.g. [GIP15, Lemma 2.1] or [BCD11, Section 2.6]. However, the product structure of the $C_u^{\gamma_1} C_v^{\gamma_2}$ -norm offers greater flexibility in regularity parameters.

Proposition 2.5 (Bilinear estimates). *Let $\alpha_j, \beta_j \in \mathbb{R} \setminus \{0\}$ and $\gamma_j \in \mathbb{R}$ be regularities, where $j = 1, 2$, and let $f, g: \mathbb{R}_{u,v}^2 \rightarrow \mathbb{R}$. Then, the following estimates hold:*

(i) (General case): If

$$\gamma_j \leq \min(\alpha_j, \beta_j) \quad \text{and} \quad \alpha_j + \beta_j > 0 \quad \text{for } j = 1, 2, \quad (2.27)$$

then we have that

$$\|fg\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.28)$$

(ii) (Low \times high-improvement): If $\alpha_1 < 0 < \beta_1$,

$$\begin{aligned} \gamma_1 &\leq \alpha_1 + (1 - \sigma)\beta_1, & \alpha_1 + \beta_1 &> 0, \\ \gamma_2 &\leq \min(\alpha_2, \beta_2), & \alpha_2 + \beta_2 &> 0, \end{aligned}$$

then

$$\|f \llcorner_\sigma^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.29)$$

(iii) (Non-resonant improvement): If

$$\begin{aligned} \gamma_1 &\leq \min(\alpha_1, \beta_1, \alpha_1 + \beta_1), \\ \gamma_2 &\leq \min(\alpha_2, \beta_2), \end{aligned} \quad \alpha_2 + \beta_2 > 0,$$

then

$$\|f \lrcorner^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.30)$$

(iv) (*Resonant improvement*): If

$$\begin{aligned}\gamma_1 &\leq \alpha_1 + \beta_1, & \alpha_1 + \beta_1 &> 0, \\ \gamma_2 &\leq \min(\alpha_2, \beta_2), & \alpha_2 + \beta_2 &> 0,\end{aligned}$$

then

$$\|f \odot^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.31)$$

Remark 2.6. The restriction $\alpha_j, \beta_j \neq 0$ is only imposed to avoid logarithmic corrections to our estimates. In (ii), the improvement lies in the upper bound on γ_1 , which is weaker than an upper bound by $\min(\alpha_1, \beta_1)$. In (iii), the improvement lies in the absence of the condition $\alpha_1 + \beta_1 > 0$. In (iv), the improvement again lies in the upper bound on γ_1 , which is similar to (ii).

Proof of Proposition 2.5: We start with the general case (i). Using a Littlewood-Paley decomposition, we obtain that

$$\begin{aligned}\|fg\|_{C_u^{\gamma_1} C_v^{\gamma_2}} &= \sup_{K_1, K_2} \left(\prod_{a=1,2} K_a^{\gamma_a} \right) \|P_{K_1}^u P_{K_2}^v (fg)\|_{L_{u,v}^\infty} \\ &\leq \sup_{K_1, K_2} \sum_{\substack{M_1, M_2, \\ N_1, N_2}} \left(\prod_{a=1,2} K_a^{\gamma_a} \right) \|P_{K_1}^u P_{K_2}^v (P_{M_1}^u P_{M_2}^v f \cdot P_{N_1}^u P_{N_2}^v g)\|_{L_{u,v}^\infty} \\ &\leq \prod_{a=1,2} \left[\sup_{K_a} \sum_{M_a, N_a} \left(1\{K_a \sim M_a \gg N_a\} + 1\{K_a \sim N_a \gg M_a\} + 1\{M_a \sim N_a \gtrsim K_a\} \right) \right. \\ &\quad \left. \times K_a^{\gamma_a} M_a^{-\alpha_a} N_a^{-\beta_a} \right] \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}.\end{aligned}$$

Thus, it suffices to prove the estimate

$$\sup_K \sum_{M, N} \left(1\{K \sim M \gg N\} + 1\{K \sim N \gg M\} + 1\{M \sim N \gtrsim K\} \right) K^\gamma M^{-\alpha} N^{-\beta} \lesssim 1 \quad (2.32)$$

for all $\alpha, \beta \in \mathbb{R} \setminus \{0\}$ and $\gamma \in \mathbb{R}$ satisfying $\gamma \leq \min(\alpha, \beta)$ and $\alpha + \beta > 0$. In the following, we write $x_- := \min(x, 0)$. We separate the proof into three sub-cases.

(a): *The high \times low-interaction* $K \sim M \gg N$. We first note that $\gamma \leq \min(\alpha, \beta)$ and $\alpha + \beta > 0$ imply that $\gamma \leq \alpha + \beta_-$. Using only that $\gamma \leq \alpha + \beta_-$, we obtain

$$\begin{aligned}\sup_K \sum_{M, N} 1\{K \sim M \gg N\} K^\gamma M^{-\alpha} N^{-\beta} &= \sup_K K^{\gamma-\alpha} \sum_N 1\{N \ll K\} N^{-\beta} \\ &\lesssim \sup_K K^{\gamma-\alpha-\beta_-} = 1.\end{aligned}$$

(b): *The low \times high-interaction* $K \sim N \gg M$. The argument is similar as in case (a) and only requires that $\gamma \leq \alpha_- + \beta$.

(c): *The high \times high-interaction* $M \sim N \gtrsim K$. Using only the condition $\alpha + \beta > 0$, we obtain

$$\sup_K \sum_{M, N} 1\{M \sim N \gtrsim K\} K^\gamma M^{-\alpha} N^{-\beta} \lesssim \sup_K K^\gamma \sum_M 1\{M \gtrsim K\} M^{-\alpha-\beta} \lesssim \sup_K K^{\gamma-\alpha-\beta}.$$

From $\alpha + \beta > 0$, it follows that $\alpha + \beta = \max(\alpha, \beta) + \min(\alpha, \beta) \geq \min(\alpha, \beta)$. Thus, the bound by one follows from the assumption $\gamma \leq \min(\alpha, \beta)$. This completes the proof of (i).

The arguments for (ii), (iii), and (iv) only require minor modifications. In (ii), the frequency-restriction in case (a) is replaced by $K \sim M \gg N \geq M^{1-\sigma}$. Instead of $\gamma \leq \alpha = \alpha + \beta_-$, we therefore only need the weaker condition $\gamma \leq \alpha + (1 - \sigma)\beta$. In (iii), the high-high interaction has been removed, which means that case (c) no longer appears. As a result, the condition $\alpha + \beta > 0$ is not needed. The new restriction on γ is a result of the identity

$$\min(\alpha + \beta_-, \alpha_- + \beta) = \min(\alpha, \beta, \alpha + \beta).$$

Finally, in (iv), we only encounter case (c), which requires that $\gamma_1 \leq \alpha_1 + \beta_1$ and $\alpha_1 + \beta_1 > 0$. \square

We now present a corollary of the proof of Proposition 2.5, which addresses families of functions.

Corollary 2.7. (Bilinear estimate for families of functions) *Let $\alpha_j, \beta_j \in \mathbb{R} \setminus \{0\}$ and $\gamma_j \in \mathbb{R}$ be regularities, where $j = 1, 2$, and assume that*

$$\gamma_j \leq \min(\alpha_j, \beta_j) \quad \text{and} \quad \alpha_j + \beta_j > 0 \quad \text{for} \quad j = 1, 2.$$

Let $(f_M)_M \subseteq C_u^{\alpha_1} C_v^{\alpha_2}$ and $(g_M)_M \subseteq C_u^{\beta_1} C_v^{\beta_2}$ be two families of functions and assume that g_M is supported on frequencies $\sim M$ in the u -variable. Then, it holds that

$$\left\| \sum_M f_M g_M \right\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \sup_M \|f_M\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \sup_M \|g_M\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.33)$$

The most important aspect of (2.33) is that the right-hand side contains a supremum and not a sum over M .

Proof. We only sketch the necessary modifications in the proof of Proposition 2.5. We first decompose

$$\sum_M f_M g_M = \sum_M f_M \otimes^u g_M + \sum_M f_M \odot^u g_M + \sum_M f_M \otimes^u g_M. \quad (2.34)$$

For the low \times high-term in (2.34), we have from frequency-support considerations that

$$\left\| \sum_M f_M \otimes^u g_M \right\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \sup_M \|f_M \otimes^u g_M\|_{C_u^{\gamma_1} C_v^{\gamma_2}}. \quad (2.35)$$

The right-hand side of (2.35) can then be estimated as before. Arguing as in the proof of Proposition 2.5, the high \times high and low \times high-terms in (2.34) can be estimated by

$$\left\| f_M \odot^u g_M \right\|_{C_u^{\gamma_1} C_v^{\gamma_2}} + \left\| f_M \otimes^u g_M \right\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim M^{\max(0, \gamma_1) - \alpha_1 - \beta_1} \|f_M\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g_M\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.36)$$

Since $\alpha_1 + \beta_1 > 0$ and

$$\gamma_1 - \alpha_1 - \beta_1 = \gamma_1 - \min(\alpha_1, \beta_1) - \max(\alpha_1, \beta_1) \leq -\max(\alpha_1, \beta_1) < 0,$$

the exponent of M in (2.36) is negative, and the contributions are therefore summable in M . \square

The following estimate is a special case of the bilinear estimate (2.28). Due to our frequent use of this estimate, however, we isolate it in the following corollary.

Corollary 2.8 (Multiplication estimate). *Let s, r' , and r be as in (2.2) and (2.5) and let $f, g: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}$. Then, it holds that*

$$\|fg\|_{C_u^{r-1}C_v^{r-1}} \lesssim \|f\|_{C_u^{1-r'}C_v^{1-r'}} \|g\|_{C_u^{r-1}C_v^{r-1}} \lesssim \|f\|_{C_u^sC_v^s} \|g\|_{C_u^{r-1}C_v^{r-1}}. \quad (2.37)$$

The inequality (2.37) is called a multiplication estimate since the map $g \mapsto f \cdot g$ is regarded as the multiplication of g with the “smooth” function f .

Proof. This follows from the bilinear estimate (2.28), $s, r', r \in (0, 1)$, $r' < r$, and $1 - r' < s$. \square

In Proposition 2.5 and its variants, we have obtained estimates for bilinear products. However, the full nonlinearity is given by $\Pi_{ij}^k(\phi)\partial_u\phi^i\partial_v\phi^j$ and therefore contains the composition of ϕ with a nonlinear function. To treat this composition, we require the following version of Bony’s para-linearization in the $C_u^{\gamma_1}C_v^{\gamma_2}$ -spaces.

Lemma 2.9 (Bony’s para-linearization). *Let $F \in C_b^\infty(\mathbb{R}^D \rightarrow \mathbb{R})$, let $0 < \gamma < 1$, and let $\phi \in C_u^\gamma C_v^\gamma$. Then, there exists a constant*

$$C_{\gamma,F} = C(\gamma, \|F\|_{L^\infty}, \dots, \|\nabla^{10}F\|_{L^\infty})$$

such that the following two properties hold:

(i) (Composition estimate): *It holds that*

$$\|F(\phi)\|_{C_u^\gamma C_v^\gamma} \leq C_{\gamma,F} (1 + \|\phi\|_{C_u^\gamma C_v^\gamma})^9 \|\phi\|_{C_u^\gamma C_v^\gamma}. \quad (2.38)$$

(ii) (Para-linearization in one variable): *For all $N > 1$, it holds that*

$$\|P_N^u(F(\phi)) - \langle \nabla F(\phi), P_N^u \phi \rangle\|_{C_u^\gamma C_v^\gamma} \leq C_{\gamma,F} N^{-\gamma} (1 + \|\phi\|_{C_u^\gamma C_v^\gamma})^9 \|\phi\|_{C_u^\gamma C_v^\gamma}, \quad (2.39)$$

$$\|P_N^v(F(\phi)) - \langle \nabla F(\phi), P_N^v \phi \rangle\|_{C_u^\gamma C_v^\gamma} \leq C_{\gamma,F} N^{-\gamma} (1 + \|\phi\|_{C_u^\gamma C_v^\gamma})^9 \|\phi\|_{C_u^\gamma C_v^\gamma}. \quad (2.40)$$

The exponents in (2.38), (2.39), and (2.40) are generously large and the exact value is irrelevant for the rest of the paper. We emphasize that (ii) only contains estimates for P_N^u and P_N^v but not the combined operator $P_{N_1}^u P_{N_2}^v$. The reason is that the size of the error term would only be bounded by inverse powers of $\min(N_1, N_2)$ instead of $\max(N_1, N_2)$, which is not sufficient for our purposes.

Proof sketch: We only sketch the argument, which is a minor modification of the Taylor expansion around low-frequencies used in the standard Besov-space version (see e.g. [BCD11, Theorem 2.89 and 2.92]). To prove (2.38) and (2.39), the simplest approach is to perform the same steps as in the proof of [BCD11, Theorem 2.89 and 2.92] in the u -variable and then apply the known results [BCD11, Theorem 2.89 and 2.92] in the remaining v -variable. To convince the reader that this reduction to the single-variable case is possible, we note that

$$\begin{aligned} \|\psi\|_{C_u^\gamma C_v^\gamma} &= \sup_{N_1, N_2 \geq 1} N_1^\gamma N_2^\gamma \|P_{N_1}^u P_{N_2}^v \psi\|_{L_{u,v}^\infty} \\ &= \sup_{N_1 \geq 1} N_1^\gamma \sup_{u \in \mathbb{R}} \|P_{N_1}^u \psi(u, v)\|_{C_v^\gamma}. \end{aligned}$$

The para-linearization (2.40) in the v -variable follows by reversing the roles of u and v . \square

2.3. Commutator estimates. In this section, we present several commutator estimates. Similar commutator estimates were already heavily used in the para-controlled approach to rough ODEs and singular parabolic SPDEs in [GIP15]. Our setting requires commutator estimates in the Bourgain-type spaces $C_u^{\gamma_1} C_v^{\gamma_2}$ instead of the standard Besov spaces, which only requires minor modifications.

Lemma 2.10. (Frequency-localized commutator estimates) *Let $\alpha_j, \beta_j \in \mathbb{R} \setminus \{0\}$ and $\gamma_j \in \mathbb{R}$, where $j = 1, 2$, be regularity parameters which satisfy $0 < \alpha_1, \beta_1, \gamma_1 < 1$ and (2.27). Furthermore, let $K \geq 1$ and let $f, g: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}$. Then, we have that*

$$\| [P_K^u, f]g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1 - \alpha_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}, \quad (2.41)$$

$$\| [P_K^u, f]P_{\lesssim K}^u g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1 - \alpha_1 - \beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}, \quad (2.42)$$

$$\| P_K^u(fg) - P_{\ll K}^u f P_K^u g - P_K^u f P_{\ll K}^u g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1 - \alpha_1 - \beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}, \quad (2.43)$$

$$\| P_K^u(fg) - f P_K^u g - P_K^u f g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1 - \alpha_1 - \beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.44)$$

Since all estimates only involve Littlewood-Paley operators in the u -variable, the proof is exactly as in the standard Besov spaces. We only present the argument for the sake of completeness.

Proof. We start by proving the estimate

$$\| [P_K^u, f]P_L^u g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1 - \alpha_1} L^{-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}} \quad (2.45)$$

for all frequency-scales K and L . The estimates (2.41) and (2.42) then follow by summing over $L \geq 1$ and $L \gtrsim K$, respectively. To prove (2.45), we distinguish the cases $L \not\sim K$ and $L \sim K$.

If $L \not\sim K$, we decompose

$$[P_K^u, f]P_L^u g = P_K^u(f P_L^u g) = P_K^u(P_{\gtrsim K}^u f P_L^u g).$$

Then, we have the estimate

$$\begin{aligned} \| P_K^u(P_{\gtrsim K}^u f P_L^u g) \|_{C_u^{\gamma_1} C_v^{\gamma_2}} &\lesssim K^{\gamma_1} \|P_{\gtrsim K}^u f P_L^u g\|_{L_u^\infty C_v^{\gamma_2}} \\ &\lesssim K^{\gamma_1 - \alpha_1} L^{-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \end{aligned} \quad (2.46)$$

If $L \sim K$, we decompose

$$\begin{aligned} [P_K^u, f]P_L^u g &= [P_K^u, P_{\lesssim K}^u f]P_L^u g + [P_K^u, P_{\gg K}^u f]P_L^u g \\ &= [P_K^u, P_{\lesssim K}^u f]P_L^u g - P_{\gg K}^u f P_K^u P_L^u g. \end{aligned} \quad (2.47)$$

Using standard commutator estimates (see e.g. [BCD11, Lemma 2.97]), the first term in (2.47) can be estimated by

$$\| [P_K^u, P_{\lesssim K}^u f]P_L^u g \|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1} \| [P_K^u, P_{\lesssim K}^u f]P_L^u g \|_{L_u^\infty C_v^{\gamma_2}}$$

$$\begin{aligned}
&\lesssim K^{\gamma_1-1} \|\partial_u P_K^u f\|_{L_u^\infty C_v^{\alpha_2}} \|P_L^u g\|_{L_u^\infty C_v^{\beta_2}} \\
&\lesssim K^{\gamma_1-\alpha_1} L^{-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}.
\end{aligned}$$

For the second term in (2.47), we estimate

$$\begin{aligned}
\|P_{\gg K}^u f P_K^u P_L^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} &\lesssim \|P_{\gg K}^u f\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \|P_K^u P_L^u g\|_{L_u^\infty C_v^{\gamma_2}} \\
&\lesssim K^{\gamma_1-\alpha_1} L^{-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}.
\end{aligned}$$

This completes the proof of (2.45). It remains to prove (2.43) and (2.44). To prove (2.43), we decompose

$$\begin{aligned}
&P_K^u(fg) - P_{\ll K}^u f P_K^u g - P_K^u f P_{\ll K}^u g \\
&= P_K^u(P_{\ll K}^u f P_{\gtrsim K}^u g + P_{\gtrsim K}^u f P_{\ll K}^u g + P_{\gtrsim K}^u f P_{\gtrsim K}^u g) - P_{\ll K}^u f P_K^u g - P_K^u f P_{\ll K}^u g \\
&= [P_K^u, P_{\ll K}^u f] P_{\gtrsim K}^u g + [P_K^u, P_{\ll K}^u g] P_{\gtrsim K}^u f + P_K^u(P_{\gtrsim K}^u f P_{\gtrsim K}^u g). \tag{2.48}
\end{aligned}$$

The first two summands in (2.48) can be bounded using (2.42). The third summand in (2.48) can be bounded using (2.46). In order to obtain the last estimate (2.44) from (2.43), it remains to prove

$$\|P_{\gtrsim K}^u f P_K^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} + \|P_K^u f P_{\gtrsim K}^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim K^{\gamma_1-\alpha_1-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \tag{2.49}$$

By symmetry, it suffices to treat the first summand in (2.49). Using frequency-support considerations, we have that

$$\|P_{\gtrsim K}^u f P_K^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} = \left\| \sum_{L \gtrsim K} P_L^u f P_K^u g \right\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim \sup_{L \gtrsim K} L^{\gamma_1} \|P_L^u f P_K^u g\|_{L_u^\infty C_v^{\gamma_2}}. \tag{2.50}$$

Using $\gamma_1 \leq \alpha_1$, the right-hand side of (2.50) can be bounded as before. \square

In the following, we state additional commutator estimates directly in terms of the para-product operators, which essentially follow from Lemma 2.10. The first corollary will be used in the PDE-analysis (Sect. 4-7) and the second corollary will be used in the ODE-analysis (Sect. 8).

In the following corollary, we prove a commutator estimate for the para-product operator \otimes_σ^u .

Corollary 2.11 (Commutators for PDE-analysis). *Let s, r' , and r be as in (2.2) and (2.5). Then, we have for all $f, g, h \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$ that*

$$\|f(g \otimes_\sigma^u h) - (fg) \otimes_\sigma^u h\|_{C_u^{r-1} C_v^{r-1}} \lesssim \|f\|_{C_u^s C_v^{1-r'}} \|g\|_{C_u^{1-r'} C_v^{r-1}} \|h\|_{C_u^{s-1+\sigma} C_v^{1-r'}}. \tag{2.51}$$

Proof. We first rewrite the argument on the left-hand side of (2.51) as

$$f(g \otimes_\sigma^u h) - (fg) \otimes_\sigma^u h = \sum_M \left(f P_{\leq M^{1-\sigma}}^u P_M^u h - P_{\leq M^{1-\sigma}}(fg) P_M^u h \right)$$

$$\begin{aligned}
&= \sum_M [f, P_{\leq M^{1-\sigma}}^u](g) P_M^u h \\
&= \sum_M [P_{> M^{1-\sigma}}^u, f](g) P_M^u h.
\end{aligned}$$

Using the bilinear estimate (Proposition 2.5) and (2.41), we have that

$$\begin{aligned}
&\left\| \sum_M [P_{> M^{1-\sigma}}^u, f](g) P_M^u h \right\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \sum_M \left\| [P_{> M^{1-\sigma}}^u, f](g) \right\|_{C_u^{1-r'} C_v^{r-1}} \left\| P_M^u h \right\|_{C_u^{r-1} C_v^{1-r'}} \\
&\lesssim \|f\|_{C_u^s C_v^{1-r'}} \|g\|_{C_u^{1-r'} C_v^{r-1}} \sum_M M^{(1-\sigma)(1-r'-s)} \left\| P_M^u h \right\|_{C_u^{r-1} C_v^{1-r'}}.
\end{aligned}$$

Since

$$(1-\sigma)(1-r'-s) + r-1 = s-1+1-2s+\eta+\sigma(r'+s-1) < s-1+\sigma,$$

this yields the desired estimate. \square

We now define three different commutator terms, which will be used in the ODE-analysis.

Definition 2.12. (Commutator terms for ODE-analysis) For all $f, g, h \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$, we define

$$\text{Com}_{\otimes}^v(f, g, h) := f(g \otimes^v h) - (fg) \otimes^v h - g(f \otimes^v h), \quad (2.52)$$

$$\text{Com}_{\otimes, \sim}^v(f, g, h) := (f \otimes^v g) \sim^v h - f(g \sim^v h), \quad (2.53)$$

$$\text{Com}_{\sim}^v(f, g, h) := (fg) \sim^v h - f(g \sim^v h) - g(f \sim^v h). \quad (2.54)$$

Equipped with Definition 2.12, we can now state the commutator estimates for the ODE-analysis.

Lemma 2.13. (Commutators for ODE-analysis) For all $f, g, h \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$, we have that

$$\left\| \text{Com}_{\otimes}^v(f, g, h) \right\|_{C_u^s C_v^{r-1}} \lesssim \|f\|_{C_u^s C_v^s} \|g\|_{C_u^s C_v^s} \|h\|_{C_u^s C_v^{r-1-s+2\eta}}, \quad (2.55)$$

$$\left\| \text{Com}_{\otimes, \sim}^v(f, g, h) \right\|_{C_u^s C_v^{r-1}} \lesssim \|f\|_{C_u^s C_v^s} \|g\|_{C_u^s C_v^s} \|h\|_{C_u^s C_v^{-1+2\eta}}, \quad (2.56)$$

$$\left\| \text{Com}_{\sim}^v(f, g, h) \right\|_{C_u^s C_v^{r-1}} \lesssim \|f\|_{C_u^s C_v^s} \|g\|_{C_u^s C_v^s} \|h\|_{C_u^s C_v^{-1+2\eta}}. \quad (2.57)$$

Proof. The three estimates can be derived from Lemma 2.10 and we omit the standard details. Similar estimates can be found in [GIP15, Lemma 2.4] and [GIP15, Lemma 2.8]. \square

We end this section with an estimate for para-products of frequency-localized functions. Strictly speaking, it is not a commutator estimate, but it has a similar flavor. This lemma will be used heavily in the analysis below, since it allows us to freely switch back and forth between para-products and Littlewood-Paley decompositions.

Lemma 2.14. (Para-products of frequency-localized functions) *Let $\alpha_j, \beta_j \in \mathbb{R} \setminus \{0\}$ and $\gamma_j \in \mathbb{R}$ be regularities, where $j = 1, 2$, and let $f, g \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$. Furthermore, let M and N be frequency scales.*

(i): *If $\gamma_2 \leq \min(\alpha_2, \beta_2)$ and $\alpha_2 + \beta_2 > 0$, then it holds that*

$$\|f \otimes_\sigma^u \tilde{P}_M^u g - P_{\leq M^{1-\sigma}}^u f \tilde{P}_M^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim M^{\gamma_1 - (1-\sigma)\alpha_1 - \beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.58)$$

(ii): *If $\gamma_1 \leq \min(\alpha_1, \beta_1)$ and $\alpha_1 + \beta_1 > 0$, then it holds that*

$$\|f \odot^v P_N^v g - \tilde{P}_N^v f P_N^v g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim N^{\gamma_2 - \alpha_2 - \beta_2} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.59)$$

Proof. We start by proving (2.58). To this end, we decompose

$$\begin{aligned} f \otimes_\sigma^u \tilde{P}_M^u g - P_{\leq M^{1-\sigma}}^u f \tilde{P}_M^u g &= \sum_{\substack{K, L: \\ K \leq L^{1-\sigma}}} P_K^u f P_L^u \tilde{P}_M^u g - \sum_{\substack{K, L: \\ K \leq M^{1-\sigma}}} P_K^u f P_L^u \tilde{P}_M^u g \\ &= \sum_{K, L} (1\{K \leq L^{1-\sigma}\} - 1\{K \leq M^{1-\sigma}\}) P_K^u f P_L^u \tilde{P}_M^u g. \end{aligned}$$

The dyadic sum is supported on frequency-scales $K \sim M^{1-\sigma}$ and $L \sim M$. For all such dyadic scales, it holds that

$$\begin{aligned} \|P_K^u f P_L^u \tilde{P}_M^u g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} &\lesssim M^{\gamma_1} \|P_K^u f P_L^u \tilde{P}_M^u g\|_{L_u^\infty C_v^{\gamma_2}} \\ &\lesssim M^{\gamma_1} \|P_K^u f\|_{L_u^\infty C_v^{\alpha_2}} \|P_L^u \tilde{P}_M^u g\|_{L_u^\infty C_v^{\beta_2}} \\ &\lesssim M^{\gamma_1} K^{-\alpha_1} L^{-\beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}} \\ &\lesssim M^{\gamma_1 - (1-\sigma)\alpha_1 - \beta_1} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \end{aligned}$$

This completes the proof of (2.58) and we now turn to (2.59). Using linearity, we can replace f in (2.59) by $P_K^v f$, where $2^{-20} \leq K/N \leq 2^{20}$. That is, f lives at frequencies comparable to N , but with a larger constant than implicit in $K \sim N$. A similar argument as for (2.58) shows that

$$\|P_K^v f \odot^v P_N^v g - P_{\ll N}^v P_K^v f P_N^v g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim N^{\gamma_2 - \alpha_2 - \beta_2} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.60)$$

Furthermore, it is easy to see that

$$\|P_K^v f \otimes^v P_N^v g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} + \|P_{\gg N}^v P_K^v f P_N^v g\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim N^{\gamma_2 - \alpha_2 - \beta_2} \|f\|_{C_u^{\alpha_1} C_v^{\alpha_2}} \|g\|_{C_u^{\beta_1} C_v^{\beta_2}}. \quad (2.61)$$

Together with the decomposition

$$\begin{aligned} P_K^v f \odot^v P_N^v g - \tilde{P}_N^v P_K^v f P_N^v g &= (P_K^v f P_N^v g - P_K^v f \otimes^v P_N^v g - P_K^v f \otimes^v P_N^v g) \\ &\quad - (P_K^v f P_N^v g - P_{\ll N}^v P_K^v f P_N^v g - P_{\gg N}^v P_K^v f P_N^v g) \\ &= P_{\ll N}^v P_K^v f P_N^v g - P_K^v f \otimes^v P_N^v g + P_{\gg N}^v P_K^v f P_N^v g - P_K^v f \otimes^v P_N^v g, \end{aligned}$$

the estimates (2.60) and (2.61) yield (2.59). \square

2.4. Integrals and traces. In this section, our main goal is to understand the Duhamel integral of the linear wave equation in null-coordinates. As we will see in Proposition 2.23 below, the Duhamel integral consists of both integral and trace operators. Before addressing combinations of integral and trace operators, we start by considering them separately.

As is clear from the wave equation (1.20), the Duhamel integral acts on both u and v -variables. Nevertheless, we start by analyzing integral operators in a single variable.

Definition 2.15 (*Single-variable integral operator*). For all functions $f \in C_c^\infty(\mathbb{R})$, we define the integral

$$If(x) := \int_0^x dy f(y). \quad (2.62)$$

The following lemma proves that the integral I gains one derivative in Hölder spaces.

Lemma 2.16 (*Single-variable integral estimate*). For all $f \in C_b^\infty(\mathbb{R})$, $\tilde{\chi} \in C_c^\infty(\mathbb{R})$, and $\gamma \in (0, \infty)$, it holds that

$$\|I[\tilde{\chi}f]\|_{C^\gamma} \lesssim_{\tilde{\chi}, \gamma} \|f\|_{C^{\gamma-1}}. \quad (2.63)$$

Remark 2.17. The cut-off function $\tilde{\chi}$, which will later be chosen as a fattened version of χ from (2.7), should be seen as a technical crutch. Since the wave maps equation exhibits finite speed of propagation, it is always possible to reduce to a compact set.

Of course, the gain of derivatives through integration is standard. For a proof of Lemma 2.16 in the Hölder-spaces from Definition 2.2, we refer to [GIP15, Lemma A.10].

We now define the integral operators I_u and I_v , which act on the u and v -variables, respectively.

Definition 2.18 (*Partial integral operators*). For all functions $f \in C_c^\infty(\mathbb{R}_{u,v}^{1+1})$, we define the partial integral operators by

$$I_u f(u, v) := \int_0^u du' f(u', v), \quad (2.64)$$

$$I_v f(u, v) := \int_0^v dv' f(u, v'). \quad (2.65)$$

The single-variable integral estimate (Lemma 2.16) directly yields estimates for the partial integral operators in our Bourgain-type space $C_u^\alpha C_v^\beta$.

Lemma 2.19 (*Partial integral estimate*). For all $f \in C^\infty(\mathbb{R}_{u,v}^{1+1})$, $\tilde{\chi} \in C_c^\infty(\mathbb{R}_{u,v}^{1+1})$, $\alpha \in (0, \infty)$, and $\beta \in \mathbb{R}$, it holds that

$$\|I_u[\tilde{\chi}f]\|_{C_u^\alpha C_v^\beta} \lesssim_{\alpha, \beta, \tilde{\chi}} \|f\|_{C_u^{\alpha-1} C_v^\beta}. \quad (2.66)$$

If instead $\alpha \in \mathbb{R}$ and $\beta \in (0, \infty)$, then

$$\|I_v[\tilde{\chi}f]\|_{C_u^\alpha C_v^\beta} \lesssim_{\alpha, \beta, \tilde{\chi}} \|f\|_{C_u^\alpha C_v^{\beta-1}}. \quad (2.67)$$

Proof. By symmetry in u and v , it suffices to prove (2.66). This estimate directly follows from the single-variable integral estimate from Lemma 2.16 and the commutativity of I_u and the Littlewood-Paley operators P_N^v . \square

This completes the analysis of integral operators and we now turn to trace operators. As mentioned above, the trace operators are examined due to their appearance in the Duhamel integral (see Proposition 2.23).

Definition 2.20 (*Trace operators*). For any $f \in C^\infty(\mathbb{R}_{u,v}^{1+1})$, we define the trace operator Tr by

$$\text{Tr } f(x) := f(x, x).$$

Furthermore, we define the trace operators Tr_u and Tr_v by

$$\text{Tr}_u f(u, v) := f(u, u) \quad \text{and} \quad \text{Tr}_v f(u, v) := f(v, v). \quad (2.68)$$

We now analyze the mapping properties of $\text{Tr}: C_u^\alpha C_v^\beta \rightarrow C_x^\gamma$.

Lemma 2.21. (Trace estimate) *Let $\alpha, \beta \in \mathbb{R} \setminus \{0\}$, $\gamma \in \mathbb{R}$, and $f \in C_u^\alpha C_v^\beta$. Then, the estimate*

$$\|\text{Tr } f\|_{C_x^\gamma} \lesssim \|f\|_{C_u^\alpha C_v^\beta}.$$

is satisfied under either of the following two conditions:

- (i) (General) *It holds that $\gamma \leq \min(\alpha, \beta)$ and $\alpha + \beta > 0$.*
- (ii) (Non-resonant) *It holds that $\gamma \leq \min(\alpha, \beta, \alpha + \beta)$ and $P_M^u P_N^v f = 0$ for all $M \sim N$.*

Proof. Using a Littlewood-Paley decomposition, we have that

$$\begin{aligned} & \|\text{Tr } f(x)\|_{C_x^\gamma} \\ &= \sup_K K^\gamma \|(P_K^x \text{Tr } f)(x)\|_{L_x^\infty} \\ &\leq \sup_K K^\gamma \sum_{M,N} \left\| \left(P_K^x \text{Tr } (P_M^u P_N^v f) \right)(x) \right\|_{L_x^\infty} \\ &\lesssim \sup_K \sum_{M,N} \left[\left(1\{K \sim M \gg N\} + 1\{K \sim N \gg M\} + 1\{M \sim N \gtrsim K\} \right) K^\gamma \|P_M^u P_N^v f(u, v)\|_{L_u^\infty L_v^\infty} \right] \\ &\lesssim \sup_K \sum_{M,N} \left[\left(1\{K \sim M \gg N\} + 1\{K \sim N \gg M\} + 1\{M \sim N \gtrsim K\} \right) K^\gamma M^{-\alpha} N^{-\beta} \right] \|f\|_{C_u^\alpha C_v^\beta}. \end{aligned}$$

This sum has been previously estimated in (2.32), and hence this completes the proof in case (i). In case (ii), which no longer requires $\alpha + \beta > 0$, the estimate follows from the argument leading to Proposition 2.5.(iii). \square

Using the trace estimate, we now prove that the $C_u^\alpha C_v^\beta$ -norms control the usual $C_t^0 C_x^\gamma$ and $C_t^1 C_x^{\gamma-1}$ -norms in Cartesian coordinates.

Corollary 2.22. (From null to Cartesian coordinates) *Let $f \in C_b^\infty(\mathbb{R}_{t,x}^{1+1})$ and $\tilde{f} \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$ satisfy*

$$f(t, x) = \tilde{f}(x - t, x + t). \quad (2.69)$$

Furthermore, let $\alpha, \beta \in \mathbb{R} \setminus \{0\}$ and let $\gamma \in \mathbb{R}$. Then, the estimate

$$\|f\|_{C_t^0 C_x^\gamma} + \|f\|_{C_t^1 C_x^{\gamma-1}} \lesssim \|\tilde{f}\|_{C_u^\alpha C_v^\beta}. \quad (2.70)$$

is satisfied under either of the following two conditions:

(i) (General) It holds that $\gamma \leq \min(\alpha, \beta)$ and $\alpha + \beta > 1$.

(ii) (Non-resonant) It holds that $\gamma \leq \min(\alpha, \beta, \alpha + \beta)$ and $P_M^u P_N^v \tilde{f} = 0$ for all $M \sim N$.

Proof. In order to utilize our trace estimate (Lemma 2.21), we define the shift operator Λ_t by

$$(\Lambda_t \tilde{f})(u, v) = \tilde{f}(u - t, v + t). \quad (2.71)$$

Due to the translation invariance of the Littlewood-Paley operators, Λ_t preserves the $C_u^\alpha C_v^\beta$ -norm. From (2.69) and (2.71), it follows that

$$f(t, x) = (\text{Tr } \Lambda_t \tilde{f})(x) \quad \text{and} \quad \partial_t f(t, x) = (\text{Tr } \partial_v \Lambda_t \tilde{f})(x) - (\text{Tr } \partial_u \Lambda_t \tilde{f})(x).$$

Using Lemma 2.21, we obtain under either of the two conditions (i) or (ii) that

$$\begin{aligned} \|f\|_{C_t^0 C_x^\gamma} + \|f\|_{C_t^1 C_x^{\gamma-1}} &\lesssim \sup_{t \in \mathbb{R}} \left(\|\text{Tr } \Lambda_t \tilde{f}\|_{C_x^\gamma} + \|\text{Tr } \partial_u \Lambda_t \tilde{f}\|_{C_x^{\gamma-1}} + \|\text{Tr } \partial_v \Lambda_t \tilde{f}\|_{C_x^{\gamma-1}} \right) \\ &\lesssim \sup_{t \in \mathbb{R}} \left(\|\Lambda_t \tilde{f}\|_{C_u^\alpha C_v^\beta} + \|\partial_u \Lambda_t \tilde{f}\|_{C_u^{\alpha-1} C_v^\beta} + \|\partial_v \Lambda_t \tilde{f}\|_{C_u^\alpha C_v^{\beta-1}} \right) \\ &\lesssim \|\tilde{f}\|_{C_u^\alpha C_v^\beta}. \end{aligned}$$

This completes the proof of (2.70). \square

Equipped with both integral and trace estimates, we now turn to the Duhamel integral.

Proposition 2.23. (Duhamel integral) *Let $F \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$ and let $\phi \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$ be a solution of the inhomogeneous linear wave equation*

$$\begin{cases} \partial_u \partial_v \phi = F & (u, v) \in \mathbb{R}^{1+1}, \\ \phi|_{u=v} = 0, \quad (\partial_v - \partial_u)\phi|_{u=v} = 0. \end{cases} \quad (2.72)$$

Then, the solution ϕ can be written as

$$\phi(u, v) = \text{Duh}[F] := - \int_u^v dv' \int_u^{v'} du' F(u', v') = - \int_u^v du' \int_{u'}^v dv' F(u', v'). \quad (2.73)$$

Equivalently, the solution ϕ is also given by

$$\phi = (I_v - \text{Tr}_u I_v)(I_u - \text{Tr}_v I_u)(F) = (I_u - \text{Tr}_v I_u)(I_v - \text{Tr}_u I_v)(F). \quad (2.74)$$

In addition, for χ^+ , χ^- defined in (2.8), we have the estimate

$$\|\chi^+ \chi^- \phi\|_{C_u^{\gamma_1} C_v^{\gamma_2}} \lesssim_{\gamma_1, \gamma_2} \|F\|_{C_u^{\gamma_1-1} C_v^{\gamma_2-1}} \quad (2.75)$$

for all $\gamma_1, \gamma_2 \in (0, 1)$ satisfying $\gamma_1 + \gamma_2 > 1$.

Proof. The identities (2.73) and (2.74) follow directly from the fundamental theorem of calculus. In order to prove the estimate (2.75), we let $\tilde{\chi} \in C_c^\infty(\mathbb{R})$ be a fattened version of χ and define

$$\tilde{\chi}^+(u, v) := \tilde{\chi}(u) \quad \text{and} \quad \tilde{\chi}^-(u, v) := \tilde{\chi}(v).$$

Due to the explicit representation (2.73), it holds that

$$\chi^+ \chi^- \phi = \chi^+ \chi^- \text{Duh}[F] = \chi^+ \chi^- \text{Duh}[\tilde{\chi}^+ \tilde{\chi}^- F].$$

Using this observation, the estimate (2.75) follows from the integral estimates (Lemma 2.19) and the trace estimate (Lemma 2.21). The condition $\gamma_1 + \gamma_2 > 1$ is needed in order to use the trace estimate for $I_u(\tilde{\chi} F)$. \square

We finish this section on integrals with a commutator estimate. To this end, we first make the following definition.

Definition 2.24. (*Integral commutator*) For all $f, g \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$, we define

$$\text{Com}_{\chi, \otimes, I}^v(f, g) := \chi^+ \chi^- \left(I_v[f \otimes^v \partial_v g](u, v) - I_v[f \otimes^v \partial_v g](u, u) - f \otimes^v g \right). \quad (2.76)$$

Lemma 2.25. (*Integrals of low \times high-terms*) Let $\zeta_{N,n} : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ be a family of functions, let $\phi^- : \mathbb{R}_v \rightarrow \mathbb{R}^D$, and let $\phi_N^- := P_N^v \phi^-$. Then, it holds that

$$\left\| \sum_N \text{Com}_{\chi, \otimes, I}^v(\zeta_{N,n}, \phi_N^{-,n}) \right\|_{C_u^s C_v^r} \lesssim \sup_N (\|\zeta_N\|_{C_u^s C_v^s}) \|\phi^-\|_{C_v^s}.$$

A similar estimate is contained in [GIP15, Lemma B.2].

Proof. We first rewrite the expression on the right-hand side of (2.76). Using the product formula for ∂_v , we obtain that

$$\begin{aligned} & I_v[\zeta_{N,n} \otimes^v \partial_v \phi_N^{-,n}](u, v) - I_v[\zeta_{N,n} \otimes^v \partial_v \phi_N^{-,n}](u, u) - (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v) \\ &= I_v[\zeta_{N,n} \otimes^v \partial_v \phi_N^{-,n}](u, v') \Big|_{v'=u}^v - (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v) \\ &= I_v[\partial_v(\zeta_{N,n} \otimes^v \phi_N^{-,n})](u, v') \Big|_{v'=u}^v - I_v[\partial_v \zeta_{N,n} \otimes^v \phi_N^{-,n}](u, v') \Big|_{v'=u}^v - (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v). \end{aligned} \quad (2.77)$$

Due to the definition of the integral operator I_v , the combined contribution of the first and third summand in (2.77) equals

$$\begin{aligned} & (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v) - (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, u) - (\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v) \\ &= -(\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, u) = -\text{Tr}_u(\zeta_{N,n} \otimes^v \phi_N^{-,n})(u). \end{aligned}$$

Using the definition of $\text{Com}_{\chi, \otimes, I}^v$, we therefore obtain that

$$\begin{aligned} & \text{Com}_{\chi, \otimes, I}^v(\zeta_{N,n}, \phi_N^{-,n})(u, v) \\ &= -\chi^+ \chi^- \text{Tr}_u(\zeta_{N,n} \otimes^v \phi_N^{-,n})(u) - \chi^+ \chi^- \left(I_v[\partial_v \zeta_{N,n} \otimes^v \phi_N^{-,n}](u, v') \Big|_{v'=u}^v \right). \end{aligned} \quad (2.78)$$

We start by estimating the first summand in (2.78). Using Lemma 2.21 and Corollary 2.7, we obtain that

$$\begin{aligned} \left\| \sum_N \text{Tr}_u(\zeta_{N,n} \otimes^v \phi_N^{-,n})(u, v) \right\|_{C_u^s C_v^r} &= \left\| \sum_N \text{Tr}_u(\zeta_{N,n} \otimes^v \phi_N^{-,n})(u) \right\|_{C_u^s} \\ &\leq \left\| \sum_N \zeta_{N,n} \otimes^v \phi_N^{-,n} \right\|_{C_u^s C_v^s} \\ &\lesssim \sup_N (\|\zeta_N\|_{C_u^s C_v^s}) \|\phi^-\|_{C_v^s}. \end{aligned}$$

We now turn to the second summand in (2.78). Using Proposition 2.5.(iii), we have that

$$\left\| \partial_v \zeta_{N,n} \otimes^v \phi_N^{-,n} \right\|_{C_u^s C_v^{r-1}} \lesssim \left\| P_{\lesssim N}^v \partial_v \zeta_N \right\|_{C_u^s C_v^\eta} \left\| \phi_N^- \right\|_{C_v^{r-1}}$$

$$\lesssim N^{1-s+\eta} N^{r-1-s} \|\zeta_N\|_{C_u^s C_v^s} \|\phi^-\|_{C_v^s}.$$

Since $r - 2s \approx -1/4$, this contribution is summable in N . The desired bound then follows from Lemma 2.19. \square

2.5. Proof of Theorem 1.2.(i). Before we can proceed with the proof of Theorem 1.2.(i), we require the following time-localization lemma. This lemma will only be used in the deterministic theory, since in the random theory a scaling argument is more convenient.

Lemma 2.26. (Time-localization) *Let $\gamma'_1, \gamma_1, \gamma_2 \in (-1/2, 0)$ satisfy $\gamma'_1 < \gamma_1$, let $\tilde{\chi} \in C_c^\infty(\mathbb{R})$, and let $\tau > 0$. Then, it holds that*

$$\left\| \tilde{\chi}\left(\frac{u-v}{\tau}\right) F \right\|_{C_u^{\gamma'_1} C_v^{\gamma_2}} \lesssim \tau^{\gamma_1 - \gamma'_1} \|F\|_{C_u^{\gamma_1} C_v^{\gamma_2}}. \quad (2.79)$$

Proof of Lemma 2.26. Throughout this proof only, we define

$$\|G\|_{\mathbb{W}_u^{\gamma_1} \mathbb{W}_v^{\gamma_2}} := \sum_{K_1, K_2} K_1^{\gamma_1} K_2^{\gamma_2} \|P_{K_1}^u P_{K_2}^v G\|_{L_u^1 L_v^1}. \quad (2.80)$$

By duality,⁶ it suffices to prove for all $\gamma'_1, \gamma_1, \gamma_2 \in (0, 1/2)$ satisfying $\gamma'_1 < \gamma_1$ that

$$\left\| \tilde{\chi}\left(\frac{u-v}{\tau}\right) G \right\|_{\mathbb{W}_u^{\gamma'_1} \mathbb{W}_v^{\gamma_2}} \lesssim \tau^{\gamma_1 - \gamma'_1} \|G\|_{\mathbb{W}_u^{\gamma_1} \mathbb{W}_v^{\gamma_2}}. \quad (2.81)$$

We now write $\tilde{\chi}_\tau(\cdot) := \tilde{\chi}(\cdot/\tau)$ and decompose $\tilde{\chi}_\tau = \sum_M P_M \tilde{\chi}_\tau$. For the dyadic components, we have the pointwise estimate

$$|P_M \tilde{\chi}_\tau|(y) \lesssim 1\{M \lesssim \tau^{-1}\} M \tau \langle My \rangle^{-10} + 1\{M \gg \tau^{-1}\} M^{-10} \langle \tau^{-1} y \rangle^{-10}$$

for all $y \in \mathbb{R}$. As a result, it follows that

$$\|P_M \tilde{\chi}_\tau\|_{L_y^1} \lesssim \tau^{\gamma_1 - \gamma'_1} M^{\gamma_1 - \gamma'_1 - 1} \quad \text{and} \quad \|P_M \tilde{\chi}_\tau\|_{L_y^\infty} \lesssim \tau^{\gamma_1 - \gamma'_1} M^{\gamma_1 - \gamma'_1}. \quad (2.82)$$

Furthermore, we decompose $G = \sum_{K_1, K_2} P_{K_1}^u P_{K_2}^v G$. By inserting both decompositions into the left-hand side of (2.81), we obtain that

$$\begin{aligned} & \left\| \tilde{\chi}\left(\frac{u-v}{\tau}\right) G \right\|_{\mathbb{W}_u^{\gamma'_1} \mathbb{W}_v^{\gamma_2}} \\ & \lesssim \sum_{M, K_1, K_2} \|P_M \tilde{\chi}_\tau(u-v) P_{K_1}^u P_{K_2}^v G\|_{\mathbb{W}_u^{\gamma'_1} \mathbb{W}_v^{\gamma_2}} \\ & \lesssim \sum_{M, K_1, K_2} \max(M, K_1)^{\gamma'_1} \max(M, K_2)^{\gamma_2} \|P_M \tilde{\chi}_\tau(u-v) P_{K_1}^u P_{K_2}^v G\|_{L_u^1 L_v^1}. \end{aligned} \quad (2.83)$$

Using Hölder's inequality and (2.82), we have the estimate

$$\begin{aligned} \|P_M \tilde{\chi}_\tau(u-v) P_{K_1}^u P_{K_2}^v G\|_{L_u^1 L_v^1} & \lesssim \|P_M \tilde{\chi}_\tau\|_{L_y^\infty} \|P_{K_1}^u P_{K_2}^v G\|_{L_u^1 L_v^1} \\ & \lesssim \tau^{\gamma_1 - \gamma'_1} M^{\gamma_1 - \gamma'_1} \|P_{K_1}^u P_{K_2}^v G\|_{L_u^1 L_v^1}. \end{aligned} \quad (2.84)$$

⁶ While the dual space of L_x^∞ is not L_x^1 , one can still characterize the L_x^∞ -norm as a supremum over integrals against L_x^1 -normalized functions. This statement generalizes to our functions spaces and is sufficient for the duality argument used here.

Using Hölder's inequality, (2.82), and Bernstein's estimate, we also have that

$$\begin{aligned} \|P_M \tilde{\chi}_\tau(u-v) P_{K_1}^\mu P_{K_2}^\nu G\|_{L_u^1 L_v^1} &\lesssim \|P_M \tilde{\chi}_\tau\|_{L_y^1} \min\left(\|P_{K_1}^\mu P_{K_2}^\nu G\|_{L_u^1 L_v^\infty}, \|P_{K_1}^\mu P_{K_2}^\nu G\|_{L_v^1 L_u^\infty}\right) \\ &\lesssim \tau^{\gamma_1-\gamma_1'} M^{\gamma_1-\gamma_1'-1} \min(K_1, K_2) \|P_{K_1}^\mu P_{K_2}^\nu G\|_{L_u^1 L_v^1}. \end{aligned} \quad (2.85)$$

After combining (2.83), (2.84), and (2.85), it remains to prove that

$$\sup_{K_1, K_2} \sum_M M^{\gamma_1-\gamma_1'-1} \max(M, K_1)^{\gamma_1'} \max(M, K_2)^{\gamma_2} \min(M, K_1, K_2) K_1^{-\gamma_1} K_2^{-\gamma_2} \lesssim 1.$$

This dyadic sum estimate easily follows by distinguishing the four cases $M \lesssim K_1, K_2$, $K_1 \lesssim M \lesssim K_2$, $K_2 \lesssim M \lesssim K_1$, and $M \gtrsim K_1, K_2$. \square

Equipped with the time-localization lemma, we now turn to the proof of deterministic well-posedness at regularities $r > 1/2$. To be more precise, let us elaborate on the notion of well-posedness. Let $\Theta > 0$ and let $0 < \tau \ll_r \Theta^{-1/(r-1/2)}$. Then, we prove for all initial data $(\phi_0, \phi_1): \mathbb{R} \rightarrow T\mathcal{M}$ satisfying

$$\|\phi_0\|_{C_x^r}, \|\phi_1\|_{C_x^{r-1}} \leq \Theta \quad (2.86)$$

that there exists a (conditionally) unique solution ϕ of (WM) in

$$(C_t^0 C_x^r \times C_t^1 C_x^{r-1})([-\tau, \tau] \times \mathbb{R} \rightarrow T\mathcal{M}).$$

In addition, we prove the continuous dependence of ϕ on the initial datum (ϕ_0, ϕ_1) .

Proof of Theorem 1.2.(i). We start the argument with a few standard reductions. Throughout the proof, we can neglect the geometric constraint $\phi(t, x) \in \mathcal{M}$. It can be recovered a-posteriori from the well-posedness theory for smooth initial data and the continuous dependence on the initial data. Due to translation invariance and finite speed of propagation, it is possible to insert space-time truncations to $|t| \lesssim \tau$ and $|x| \lesssim 1$ into the equation. Finally, Corollary 2.22 and the condition $r > 1/2$ allow us to argue entirely in null coordinates. In total, these reductions lead to the fixed-point problem

$$\begin{aligned} \phi^k(u, v) &= (\chi^+ \phi^{+,k})(u) + (\chi^- \phi^{-,k})(v) - \chi^+(u) \chi^-(v) \\ &\quad \text{Duh} \left[\chi((v-u)/\tau) \Pi_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right](u, v), \end{aligned} \quad (2.87)$$

where ϕ^+ and ϕ^- are the right and left-moving linear waves, respectively. We now solve (2.87) using a contraction argument. In order to later gain a power of τ , we introduce the parameter $r' := (1/2 + r)/2$, which is between $1/2$ and r . Using this new parameter, we define the norm

$$\|\phi\|_{\mathcal{S}} = \|\phi\|_{C_u^r C_v^{r'}} + \|\phi\|_{C_u^{r'} C_v^r}.$$

We recall from (2.86) that $\Theta > 0$ denotes the size of the initial data. For a constant $C = C(r)$, which remains to be chosen, we define the ball of radius $C\Theta$ by

$$\mathcal{S}_{C\Theta} := \{\phi: \|\phi\|_{\mathcal{S}} \leq C\Theta\}.$$

Finally, we define a map Γ , which encodes the right-hand side of (2.87), by

$$(\Gamma\phi)^k(u, v) := (\chi^+ \phi^{+,k})(u) + (\chi^- \phi^{-,k})(v) - \chi^+(u) \chi^-(v)$$

$$\text{Duh} \left[\chi \left((v - u)/\tau \right) \mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right] (u, v). \quad (2.88)$$

In order to complete the proof, it remains to prove that Γ is a contraction on $\mathcal{S}_{C\Theta}$ and that the resulting fixed-point depends continuously on (ϕ_0, ϕ_1) . Since the contraction property and continuous dependence follow from similar arguments, we only prove that Γ maps $\mathcal{S}_{C\Theta}$ back into itself. To this end, we let $\phi \in \mathcal{S}_{C\Theta}$ be arbitrary and prove that $\|\Gamma\phi\|_{\mathcal{S}} \leq C\Theta$. By symmetry, it suffices to prove that

$$\|\Gamma\phi\|_{C_u^{r'} C_v^r} \leq C\Theta/2. \quad (2.89)$$

For the linear waves in (2.88), it follows directly from the definitions that

$$\|\chi^+(u)\phi^+(u)\|_{C_u^{r'} C_v^r} + \|\chi^-(v)\phi^-(v)\|_{C_u^{r'} C_v^r} \lesssim \|\phi_0\|_{C_x^r} + \|\phi_1\|_{C_x^{r-1}} \lesssim \Theta. \quad (2.90)$$

For the nonlinear term, it follows from Proposition 2.23 and Lemma 2.26 that

$$\begin{aligned} & \left\| \chi^+ \chi^- \text{Duh} \left[\chi \left((v - u)/\tau \right) \mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right] \right\|_{C_u^{r'} C_v^r} \\ & \lesssim \left\| \chi \left((v - u)/\tau \right) \mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right\|_{C_u^{r'-1} C_v^{r-1}} \\ & \lesssim \tau^{r-r'} \left\| \mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right\|_{C_u^{r-1} C_v^{r-1}}. \end{aligned} \quad (2.91)$$

Using the multiplication estimate (Corollary 2.8), the bilinear estimate (Proposition 2.5), and the composition estimate (Lemma 2.9), we obtain that

$$\begin{aligned} \left\| \mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right\|_{C_u^{r-1} C_v^{r-1}} & \lesssim \left\| \mathbb{I}_{ij}^k(\phi) \right\|_{C_u^{r'} C_v^{r'}} \left\| \partial_u \phi^i \right\|_{C_u^{r-1} C_v^{r'}} \left\| \partial_v \phi^j \right\|_{C_u^{r'} C_v^{r-1}} \\ & \lesssim (1 + C\Theta)^{10} (C\Theta)^2. \end{aligned} \quad (2.92)$$

By combining (2.90), (2.91), and (2.92), it follows that

$$\|\Gamma\phi\|_{C_u^{r'} C_v^r} \lesssim_{r,r'} \Theta + \tau^{r-r'} (1 + C\Theta)^{10} (C\Theta)^2.$$

The desired estimate (2.89) now follows by first choosing $C = C(r) \geq 1$ sufficiently large and then choosing $0 < \tau \ll_r \Theta^{-1/(r-1/2)}$ sufficiently small. \square

Before the end of this section, we state a lemma which essentially follows from the previous proof.

Lemma 2.27. (Finite speed of propagation and uniqueness in $C_u^r C_v^r$) *Let $\chi \in C^\infty(\mathbb{R})$ satisfy (2.7) and let $\tilde{\chi} \in C_c^\infty(\mathbb{R})$ satisfy $\tilde{\chi}(x) = 1$ for all $x \in [-4, 4]$. Furthermore, let $\phi^+, \phi^- : \mathbb{R}_x \rightarrow \mathbb{R}^D$ and $\phi, \tilde{\phi} : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ satisfy the following conditions:*

- (i) *The linear waves satisfy $\phi^+, \phi^- \in C_x^r$.*
- (ii) *The maps ϕ and $\tilde{\phi}$ locally belong to $C_u^r C_v^r$, i.e., $\tilde{\chi}^+ \tilde{\chi}^- \phi \in C_u^r C_v^r$ and $\tilde{\chi}^+ \tilde{\chi}^- \tilde{\phi} \in C_u^r C_v^r$.*
- (iii) *The map ϕ solves the Duhamel integral problem*

$$\phi^k(u, v) = \phi^{+,k}(u) + \phi^{-,k}(v) - \text{Duh} \left[\mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j \right].$$

- (iv) *The map $\tilde{\phi}$ solves the localized Duhamel integral problem*

$$\tilde{\phi}^k(u, v) = \phi^{+,k}(u) + \phi^{-,k}(v) - \chi^+ \chi^- \text{Duh} \left[\mathbb{I}_{ij}^k(\tilde{\phi}) \partial_u \tilde{\phi}^i \partial_v \tilde{\phi}^j \right].$$

Then, it holds that $\phi(u, v) = \tilde{\phi}(u, v)$ for all $u, v \in [-2, 2]$.

Proof. The lemma follows from the same estimates as in the proof of Theorem 1.2.(i) and a continuity argument. In the continuity argument, it is best to work with the usual localization of our norms, i.e.,

$$\|\psi\|_{C_u^r C_v^r([- \tau, \tau]^2)} := \inf\{\|\zeta\|_{C_u^r C_v^r} : \zeta(u, v) = \psi(u, v) \text{ for all } u, v \in [-\tau, \tau]^2\}.$$

We leave the standard details to the reader. \square

2.6. Brownian paths. In this section we introduce smooth approximations of the Brownian paths $B: \mathbb{R} \rightarrow \mathcal{M}$ and the white noise velocities $V \in B^*T\mathcal{M}$, and we establish quantitative approximation properties that will be needed for the proof of Theorem 1.4.

As explained in the introduction, we take an extrinsic approach to construct Brownian paths $B: \mathbb{R} \rightarrow \mathcal{M}$ on our submanifold $\mathcal{M} \hookrightarrow \mathbb{R}^D$ embedded into the ambient Euclidean space \mathbb{R}^D . To this end we denote by $W: \mathbb{R} \rightarrow \mathcal{M}$ a D -dimensional Euclidean Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, we introduce for each $p \in \mathcal{M}$ the orthogonal projection $P(p): \mathbb{R}^D \rightarrow T_p\mathcal{M}$ from \mathbb{R}^D to the tangent space $T_p\mathcal{M}$, and we fix a reference point $B_0 \in \mathcal{M}$. Following [Hsu02, Chapter 3.2] we then obtain a Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ by solving the following Stratonovich stochastic differential equation on \mathcal{M}

$$dB(x) = P(B(x)) \circ dW(x), \quad B(0) = B_0 \in \mathcal{M}, \quad (2.93)$$

which is driven by the D -dimensional Euclidean Brownian motion $W(x)$.⁷

Remark 2.28. Intrinsically, Brownian motion on a Riemannian manifold can be defined as a diffusion process generated by half of the Laplace-Beltrami operator on the manifold. We refer to [Hsu02, Chapter 3] for more background. For the extrinsic approach to obtain Brownian motion as a solution to a stochastic differential equation driven by an ambient Euclidean Brownian motion the Stratonovich formulation (2.93) is key. Indeed, only the Stratonovich formulation preserves the classical chain rule, which is essential for proving that $B(x) \in \mathcal{M}$. In contrast, the solution to the Itô stochastic differential equation

$$dX(x) = P(X(x))dW(x), \quad X(0) = X_0 \in \mathcal{M},$$

does not map into the manifold \mathcal{M} .

In order to define suitable smooth approximations of the Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$, we introduce a countable family of smooth functions $W^\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^D$, $\varepsilon > 0$, that approximate the D -dimensional ambient Brownian motion W . Specifically, we set

$$W^\varepsilon := P_{\leq \varepsilon^{-1}} W = K_\varepsilon * W, \quad \varepsilon := 2^{-k}, \quad k \in \mathbb{N},$$

where $K_\varepsilon(x) := \varepsilon^{-1} \tilde{\rho}(\varepsilon^{-1}x)$. While we consider the variable ε to be restricted to dyadic numbers, we do not make this more explicit in our notation. We obtain a corresponding family of smooth functions $B^\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$, $\varepsilon > 0$, as solutions to the classical ordinary differential equations

$$\partial_x B^\varepsilon = P(B^\varepsilon) \partial_x W^\varepsilon, \quad B^\varepsilon(0) = B_0 \in \mathcal{M}. \quad (2.94)$$

⁷ While it is customary to write B_x or W_x for Brownian motions indexed by the variable x , we use the notation $B(x)$, respectively $W(x)$, which we consider to be more in line with the notation in the rest of the paper.

In view of the definition of the Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ as a solution to the Stratonovich differential equation (2.93), the family of smooth functions $(B^\varepsilon)_{\varepsilon>0}$ should provide good approximations of B as $\varepsilon \rightarrow 0$. To quantify this we work with smooth local approximations.

For any $\tau > 0$ and $x_0 \in \mathbb{R}$, we first define the re-scaled and translated smooth approximation

$$B_{\tau,x_0}^\varepsilon(x) := B^\varepsilon(\tau x + x_0) \quad (2.95)$$

and the associated re-scaled and translated approximation of the Euclidean Brownian motion

$$W_{\tau,x_0}^\varepsilon(x) := \tau^{-\frac{1}{2}}(W^\varepsilon(\tau x + x_0) - W^\varepsilon(x_0)).$$

From (2.94) it follows that $B_{\tau,x_0}^\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$ satisfies the classical ordinary differential equation

$$\partial_x B_{\tau,x_0}^\varepsilon = \tau^{\frac{1}{2}} P(B_{\tau,x_0}^\varepsilon) \partial_x W_{\tau,x_0}^\varepsilon, \quad B_{\tau,x_0}^\varepsilon(0) = B_{\tau,x_0}^\varepsilon(0).$$

Next, we introduce localized versions of the smooth approximations B_{τ,x_0}^ε . To this end, we let $\chi \in C_c^\infty(\mathbb{R})$ be as in (2.7). Then we define $B_{\tau,x_0,\text{loc}}^\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$ as the solution to the classical ordinary differential equation

$$\partial_x B_{\tau,x_0,\text{loc}}^\varepsilon = \tau^{\frac{1}{2}} P(B_{\tau,x_0,\text{loc}}^\varepsilon) \partial_x (\chi(x) W_{\tau,x_0}^\varepsilon), \quad B_{\tau,x_0,\text{loc}}^\varepsilon(0) = B_{\tau,x_0}^\varepsilon(0). \quad (2.96)$$

As a technical tool we will make use of the following moment bounds for weighted L_x^∞ estimates of the real-valued Brownian motions W_{τ,x_0}^j , $1 \leq j \leq D$.

Lemma 2.29. *Let $\tau > 0$ and $x_0 \in \mathbb{R}$ be arbitrary. For any $\sigma > \frac{1}{2}$ there exists a constant $C(\sigma) > 0$ such that for all $p \geq 2$ and for all $1 \leq j \leq D$ we have*

$$\|\langle x \rangle^{-\sigma} W_{\tau,x_0}^j\|_{L_x^\infty} \|_{L_\omega^p(\Omega)} \leq C(\sigma) \sqrt{p}. \quad (2.97)$$

Proof. By scaling and translation invariance of Brownian motion, the laws of W_{τ,x_0}^j and W^j are the same under \mathbb{P} . Correspondingly, it suffices to prove (2.97) for W^j . Using a dyadic decomposition of the line we obtain

$$\begin{aligned} \|\langle x \rangle^{-\sigma} W^j\|_{L_x^\infty} \|_{L_\omega^p} &\leq \left\| \sup_{|x| \leq 1} \langle x \rangle^{-\sigma} |W^j(x)| \right\|_{L_\omega^p} + \sum_{K \geq 1} \left\| \sup_{K \leq |x| \leq 2K} \langle x \rangle^{-\sigma} |W^j(x)| \right\|_{L_\omega^p} \\ &\lesssim \left\| \sup_{|x| \leq 1} |W^j(x)| \right\|_{L_\omega^p} + \sum_{K \geq 1} K^{-\sigma} \left\| \sup_{0 \leq |x| \leq 2K} |W^j(x)| \right\|_{L_\omega^p}. \end{aligned}$$

By Doob's maximal inequality [KS91, Theorem 1.3.8] we can bound the preceding line by

$$p' \|W^j(1)\|_{L_\omega^p} + \sum_{K \geq 1} K^{-\sigma} p' \|W^j(2K)\|_{L_\omega^p},$$

where $p' = \frac{p}{p-1} \leq 2$ denotes the conjugate exponent of $p \geq 2$. Invoking the well-known moment bounds for standard Euclidean Brownian motion

$$\|W^j(x)\|_{L_\omega^p(\Omega)} \lesssim \sqrt{p}|x|^{\frac{1}{2}},$$

we conclude uniformly for all $p \geq 2$

$$\|\langle x \rangle^{-\sigma} W^j\|_{L_x^\infty} \|_{L_\omega^p} \lesssim p' \sqrt{p} \left(1 + \sum_{K \geq 1} K^{-\sigma} K^{\frac{1}{2}}\right) \lesssim \sigma \sqrt{p},$$

which finishes the proof of the lemma. \square

Next, we establish $\text{high} \times \text{high} \rightarrow \text{low}$ -bounds for Euclidean Brownian motions.

Lemma 2.30. (*High \times high \rightarrow low-bounds for Euclidean Brownian motion*) *Let $\tau > 0$, $x_0 \in \mathbb{R}$, and $0 < \alpha < \frac{1}{2}$ be arbitrary. For any $\lambda > 0$ there exists an event $\tilde{\mathcal{E}}_\lambda(\tau, x_0) \subseteq \Omega$ satisfying*

$$\mathbb{P}(\tilde{\mathcal{E}}_\lambda(\tau, x_0)) \geq 1 - c^{-1} \exp(-c\lambda) \quad (2.98)$$

and on which the following two estimates hold:

(i) *Uniform high \times high \rightarrow low-bounds: For all $1 \leq j_1, j_2 \leq D$, it holds that*

$$\sup_{\varepsilon > 0} \sup_{M \sim N} M^\alpha \|\chi(\cdot)^2 P_M W_{\tau, x_0}^{\varepsilon, j_1} \partial_x P_N W_{\tau, x_0}^{\varepsilon, j_2}\|_{C^{\alpha-1}} \leq \lambda. \quad (2.99)$$

(ii) *Convergence of high \times high \rightarrow low-term: For all $1 \leq j_1, j_2 \leq D$, it holds that*

$$\lim_{\varepsilon \rightarrow 0} \sup_{M \sim N} M^\alpha \|\chi(\cdot)^2 (P_M W_{\tau, x_0}^{\varepsilon, j_1} \partial_x P_N W_{\tau, x_0}^{\varepsilon, j_2} - P_M W_{\tau, x_0}^{j_1} \partial_x P_N W_{\tau, x_0}^{j_2})\|_{C^{\alpha-1}} = 0. \quad (2.100)$$

Proof. We begin with the proof of the uniform $\text{high} \times \text{high} \rightarrow \text{low}$ bounds (2.99). Fix $1 \leq j_1, j_2 \leq D$. We consider the random variable

$$X := \sup_{\varepsilon > 0} \sup_{M \sim N} M^\alpha \|\chi(\cdot)^2 P_M W^{\varepsilon, j_1} \partial_x P_N W^{\varepsilon, j_2}\|_{C^{\alpha-1}}, \quad (2.101)$$

and define the event $\tilde{\mathcal{E}}_\lambda(\tau, x_0) := \{X \leq \lambda\}$. For all $p \geq 2$ we seek to establish the moment bounds $\|X\|_{L_\omega^p} \lesssim p$. By Chebyshev's inequality these moment bounds imply the tail estimate $\mathbb{P}(\tilde{\mathcal{E}}_\lambda(\tau, x_0)^c) \leq c^{-1} e^{-c\lambda}$ for any $\lambda > 0$ for some absolute constant $c > 0$. The main work now goes into establishing the moment bounds. We reduce their proof to a computation for random Fourier series representations of Brownian motions on unit-sized intervals. This reduction is achieved in several steps. We have

$$\|X\|_{L_\omega^p} \lesssim \sum_{M \sim N} M^\alpha \left\| \sup_{\varepsilon > 0} \|\chi(\cdot)^2 P_M W^{\varepsilon, j_1} \partial_x P_N W^{\varepsilon, j_2}\|_{C^{\alpha-1}} \right\|_{L_\omega^p}. \quad (2.102)$$

Step 1: Estimating the low frequencies $M \sim N \lesssim 1$. We first dispense of the low frequency contributions to (2.102). Using the embedding $L^\infty \hookrightarrow C^{\alpha-1}$ and Hölder's inequality, they can be estimated crudely by

$$\begin{aligned} & \sum_{M \sim N \lesssim 1} M^\alpha \left\| \sup_{\varepsilon > 0} \|\chi(\cdot)^2 P_M W^{\varepsilon, j_1} \partial_x P_N W^{\varepsilon, j_2}\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \lesssim 1} M^\alpha \sum_{L, L' \sim M} \left\| \|\chi(\cdot)^2 P_M P_L W^{j_1} \partial_x P_N P_{L'} W^{j_2}\|_{L_x^\infty} \right\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \lesssim 1} M^\alpha \sum_{L, L' \sim M} \|\chi(\cdot) P_M P_L \langle \cdot \rangle\|_{L_x^\infty \rightarrow L_x^\infty} \|\langle x \rangle^{-1} W^{j_1}\|_{L_x^\infty} \|_{L_\omega^{2p}} \\ & \quad \times \|\chi(\cdot) \partial_x P_N P_{L'} \langle \cdot \rangle\|_{L_x^\infty \rightarrow L_x^\infty} \|\langle x \rangle^{-1} W^{j_2}\|_{L_x^\infty} \|_{L_\omega^{2p}}. \end{aligned} \quad (2.103)$$

The following operator norm bounds hold uniformly for the low frequency configurations $M \sim N \lesssim 1$ and $L, L' \sim M$,

$$\|\chi(\cdot)P_M P_L \langle \cdot \rangle\|_{L_x^\infty \rightarrow L_x^\infty} + \|\chi(\cdot)\partial_x P_N P_{L'} \langle \cdot \rangle\|_{L_x^\infty \rightarrow L_x^\infty} \lesssim 1.$$

Thus, by Lemma 2.29 the last line of (2.103) can be bounded by

$$\sum_{M \sim N \lesssim 1} M^\alpha \sum_{L, L' \sim M} p \lesssim p,$$

as desired.

Step 2: Reduction to a random Fourier series computation. We begin to estimate the high-frequency contributions to (2.102). To this end we introduce a fattened version $\tilde{\chi}(x)$ of the bump function $\chi(x)$, which is as in (2.7). For notational purposes, it is convenient to require that $\tilde{\chi}(x) = 1$ for all $x \in [-5/2, 5/2]$ and that the support of $\tilde{\chi}$ is contained in $[-3, 3] \subseteq [-\pi, \pi]$. Then we decompose into

$$\begin{aligned} & \sum_{M \sim N \gg 1} M^\alpha \left\| \sup_{\varepsilon > 0} \|\chi(\cdot)^2 P_M W^{\varepsilon, j_1} \partial_x P_N W^{\varepsilon, j_2}\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L W^{j_1} \partial_x P_N P_{L'} W^{j_2}\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L (\tilde{\chi}(\cdot) W^{j_1}) \partial_x P_N P_{L'} (\tilde{\chi}(\cdot) W^{j_2})\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \quad + \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L ((1 - \tilde{\chi}(\cdot)) W^{j_1}) \partial_x P_N P_{L'} (\tilde{\chi}(\cdot) W^{j_2})\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \quad + \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L (\tilde{\chi}(\cdot) W^{j_1}) \partial_x P_N P_{L'} ((1 - \tilde{\chi}(\cdot)) W^{j_2})\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \quad + \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L ((1 - \tilde{\chi}(\cdot)) W^{j_1}) \partial_x P_N P_{L'} ((1 - \tilde{\chi}(\cdot)) W^{j_2})\|_{C^{\alpha-1}} \right\|_{L_\omega^p}. \end{aligned} \quad (2.104)$$

The first term on the right-hand side of (2.104) is the main term that will be estimated in the next step. All other terms have at least one input with a mismatched spatial support that allows for a simpler treatment. Indeed, we can bound the second term on the right-hand side of (2.104) using the embedding $L^\infty \hookrightarrow C^{\alpha-1}$ by

$$\begin{aligned} & \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L ((1 - \tilde{\chi}(\cdot)) W^{j_1}) \partial_x P_N P_{L'} (\tilde{\chi}(\cdot) W^{j_2})\|_{C^{\alpha-1}} \right\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \|\chi(\cdot)^2 P_M P_L ((1 - \tilde{\chi}(\cdot)) \langle \cdot \rangle)\|_{L_x^\infty \rightarrow L_x^\infty} \right\|_{L_\omega^p} \left\| \|\langle x \rangle^{-1} W^{j_1}\|_{L_x^\infty} \right\|_{L_\omega^{2p}} \\ & \quad \times \left\| \|\partial_x P_N P_{L'} (\tilde{\chi}(\cdot) \langle \cdot \rangle)\|_{L_x^\infty \rightarrow L_x^\infty} \right\|_{L_\omega^p} \left\| \|\langle x \rangle^{-1} W^{j_2}\|_{L_x^\infty} \right\|_{L_\omega^{2p}}. \end{aligned} \quad (2.105)$$

Due to the mismatched spatial supports of the cutoffs $\chi(\cdot)$ and $1 - \tilde{\chi}(\cdot)$, we have for all $M \sim L \gg 1$ and $A \geq 1$ that,

$$\|\chi(\cdot)^2 P_M P_L ((1 - \tilde{\chi}(\cdot)) \langle \cdot \rangle)\|_{L_x^\infty \rightarrow L_x^\infty} \lesssim_A M^{-A}.$$

By Bernstein estimates we also have uniformly for all $N \sim L' \gg 1$ that

$$\|\partial_x P_N P_{L'}(\tilde{\chi}(\cdot)(\cdot))\|_{L_x^\infty \rightarrow L_x^\infty} \lesssim N.$$

Hence, by Lemma 2.29 we can bound the last line of (2.105) as desired by

$$\sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \sum_{K \geq 1} K^{\alpha-1} M^{-2} N p \lesssim p.$$

The last two terms on the right-hand side of (2.105) can be estimated analogously.

Step 3: The main random Fourier series computation. We now turn to estimating the more delicate first term on the right-hand side of (2.104). To this end we pass to a representation of real-valued Brownian motion in terms of a random Fourier series. Recall that for any countable orthonormal basis $\{\psi_a\}_{a \in \mathbb{N}}$ of $L_x^2(\mathbb{R})$, there exists a family of independent standard Gaussian random variables $\{h_a\}_{a \in \mathbb{N}}$ such that

$$W(x) = \sum_{a=1}^{\infty} h_a \langle \mathbb{1}_{[0,x]}, \psi_a \rangle,$$

where $\langle f, g \rangle := \int_{\mathbb{R}} f(y)g(y)dy$ denotes the inner product on $L_x^2(\mathbb{R})$, see for instance [Eva13, Chapter 3.3] or [Nua06, Chapter 1]. Since the family $\{\frac{e^{imx}}{\sqrt{2\pi}}\}_{m \in \mathbb{Z}}$ forms an orthonormal basis of the (complex) vector space $L_x^2([-\pi, \pi])$, for $x \in [-\pi, \pi]$ the one-dimensional real-valued Brownian motions $W^j(x)$, $1 \leq j \leq D$, can be written as

$$\begin{aligned} W^j(x) &= \sum_{m \in \mathbb{Z}} \frac{g_m^j}{\sqrt{2\pi}} \langle \mathbb{1}_{[0,x]}(\cdot), e^{im\cdot} \rangle \\ &= \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}im} e^{imx} + \left(\frac{g_0^j}{\sqrt{2\pi}} x - \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}im} \right), \end{aligned}$$

where $\{g_m^j\}_{m \in \mathbb{Z}}$, $1 \leq j \leq D$, are families of independent standard complex-valued Gaussian random variables satisfying the constraints $g_m^j = \overline{g_{-m}^j}$. The latter constraints are necessary to ensure that $W^j(x)$ are real-valued Brownian motions. Correspondingly, we have for $x \in [-\pi, \pi]$ and $1 \leq j \leq D$ that

$$\partial_x(\tilde{\chi}(x)W^j(x)) = \tilde{\chi}(x) \left(\sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}} e^{imx} + \frac{g_0^j}{\sqrt{2\pi}} \right) + (\partial_x \tilde{\chi})(x)W^j(x).$$

Then we further decompose into

$$[P_M P_L(\tilde{\chi}W^j)](x) = \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}im} \rho_M(m) \rho_L(m) \tilde{\chi}(x) e^{imx} + \mathcal{R}_{M,L}^j(x), \quad (2.106)$$

where we introduce the remainder term

$$\begin{aligned} \mathcal{R}_{M,L}^j(x) &:= \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}im} \varphi_{M,L;m}(x) \\ &\quad + \left(\frac{g_0^j}{\sqrt{2\pi}} [P_M P_L(\tilde{\chi}(y)y)](x) - \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^j}{\sqrt{2\pi}im} [P_M P_L \tilde{\chi}](x) \right) \end{aligned}$$

with

$$\varphi_{M,L;m}(x) := [P_M P_L(\tilde{\chi}(\cdot)e^{im\cdot})](x) - \rho_M(m)\rho_L(m)\tilde{\chi}(x)e^{imx}.$$

All terms in $\mathcal{R}_{M,L}^j$ gain regularity compared with the first term on the right-hand side of (2.106) in the sense that all terms in $\mathcal{R}_{M,L}^j$ come with an additional inverse power of M . Indeed, we have uniformly for all $M \sim L \gg 1$, all $m \in \mathbb{Z} \setminus \{0\}$, and all $x \in \mathbb{R}$ that

$$|\varphi_{M,L;m}(x)| \lesssim_{A,B} \langle x \rangle^{-A} M^{-1} (1\{|m| \lesssim M\} + 1\{|m| \gg M\} \langle |m| \rangle^{-B}).$$

Moreover, by Bernstein estimates we have $\|P_M P_L(\tilde{\chi}(y)y)\|_{L_x^\infty} + \|P_M P_L \tilde{\chi}\|_{L_x^\infty} \lesssim_A M^{-A}$ uniformly for all $M \sim L \gg 1$. Similarly, we decompose

$$[P_N P_{L'} \partial_x (\tilde{\chi} W^j)](x) = \sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{g_n^j}{\sqrt{2\pi}} \rho_N(n) \rho_{L'}(n) \tilde{\chi}(x) e^{inx} + \tilde{\mathcal{R}}_{N,L'}^j(x) \quad (2.107)$$

with an analogous remainder term that gains regularity compared with the first term on the right-hand side of (2.107). We thus arrive at the decomposition

$$\begin{aligned} & \chi(x)^2 [P_M P_L(\tilde{\chi} W^{j_1})](x) [\partial_x P_N P_{L'}(\tilde{\chi} W^{j_2})](x) \\ &= \chi(x)^2 \left(\sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1}}{\sqrt{2\pi} i m} \rho_M(m) \rho_L(m) e^{imx} \right) \left(\sum_{n \in \mathbb{Z} \setminus \{0\}} \frac{g_n^{j_2}}{\sqrt{2\pi}} \rho_N(n) \rho_{L'}(n) e^{inx} \right) + \{\text{better}\}, \\ &= \sum_{m,n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) \chi(x)^2 e^{i(m+n)x} + \{\text{better}\}, \end{aligned} \quad (2.108)$$

where all product terms with at least one input at better regularity are grouped into the quantity $\{\text{better}\}$. In what follows we only estimate the contribution of the delicate first term and leave the treatment of the contributions of all other more regular terms to the reader.

Let $q \gg 1$ satisfy $\alpha + \frac{1}{q} < \frac{1}{2}$. Using Hölder's inequality in the ω -variable, it suffices to prove the moment estimate for $p \geq q$. Using Bernstein estimates and Minkowski's integral inequality, we obtain that

$$\begin{aligned} & \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} M^\alpha \left\| \chi(\cdot)^2 P_M P_L(\tilde{\chi} W^{j_1}) \partial_x P_N P_{L'}(\tilde{\chi} W^{j_2}) \right\|_{C^{\alpha-1}} \Big\|_{L_\omega^p} \\ & \lesssim \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^\alpha K^{\alpha-1} K^{\frac{1}{q}} \left\| \left\| P_K(\chi(\cdot)^2 P_M P_L(\tilde{\chi} W^{j_1}) \partial_x P_N P_{L'}(\tilde{\chi} W^{j_2})) \right\|_{L_\omega^p} \right\|_{L_x^q}. \end{aligned}$$

Inserting the first term on the right-hand side of (2.108) and using Gaussian hypercontractivity yields

$$\begin{aligned} & \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^\alpha K^{\alpha-1} K^{\frac{1}{q}} \\ & \times \left\| \left\| \sum_{m,n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \right\|_{L_\omega^p} \right\|_{L_x^q} \end{aligned}$$

$$\begin{aligned}
&\lesssim p \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^\alpha K^{\alpha-1} K^{\frac{1}{q}} \\
&\quad \times \left\| \left\| \sum_{m, n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \right\|_{L_\omega^2} \right\|_{L_x^q}.
\end{aligned} \tag{2.109}$$

We now decompose the second-order Gaussian chaos in (2.109) into a non-resonant and resonant component. To be precise, we decompose

$$\begin{aligned}
&\sum_{m, n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \\
&= \sum_{m, n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2} - \delta_{m+n=0} \delta_{j_1=j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \\
&\quad + \delta_{j_1=j_2} \sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{1}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(m) \rho_{L'}(m) P_K(\chi^2)(x).
\end{aligned} \tag{2.110}$$

$$\tag{2.111}$$

We first address the resonant term (2.111). At first sight, it may seem like this term is of order one, which would lead to a divergence in the sum over M . However, since $m \in \mathbb{Z} \setminus \{0\} \mapsto 1/m$ is odd and the cut-off function $m \in \mathbb{Z} \setminus \{0\} \mapsto \rho_M(m) \rho_L(m) \rho_N(m)$ is even, it holds that

$$\sum_{m \in \mathbb{Z} \setminus \{0\}} \frac{1}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(m) \rho_{L'}(m) = 0.$$

In particular, the resonant term (2.111) vanishes. It remains to treat the contribution of the non-resonant term (2.110). To this end, we note that

$$\left| [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \right| \lesssim_{A,B} \langle x \rangle^{-A} (1\{|m+n| \lesssim K\} + 1\{|m+n| \gg K\} \langle m+n \rangle^{-B}) \tag{2.112}$$

for all $K \geq 1$, $m, n \in \mathbb{Z}$, and $x \in \mathbb{R}$. Using L_ω^2 -orthogonality and (2.112), we obtain the bound

$$\begin{aligned}
&p \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^\alpha K^{\alpha-1} K^{\frac{1}{q}} \\
&\quad \times \left\| \left\| \sum_{m, n \in \mathbb{Z} \setminus \{0\}} \frac{g_m^{j_1} g_n^{j_2} - \delta_{m+n=0} \delta_{j_1=j_2}}{2\pi i m} \rho_M(m) \rho_L(m) \rho_N(n) \rho_{L'}(n) [P_K(\chi(\cdot)^2 e^{i(m+n)\cdot})](x) \right\|_{L_\omega^2} \right\|_{L_x^q} \\
&\lesssim p \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^{\alpha-1} K^{\alpha-1} K^{\frac{1}{q}} \times \left\| 1\{|m| \sim M\} 1\{|n| \sim N\} P_K(\chi(\cdot)^2 e^{i(m+n)\cdot}) \right\|_{\ell_{m,n}^2} \left\| \right\|_{L_x^q} \\
&\lesssim p \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^{\alpha-1} K^{\alpha-1} K^{\frac{1}{q}} \\
&\quad \times \left\| 1\{|m| \sim M\} 1\{|n| \sim N\} (1\{|m+n| \sim K\} + 1\{|m+n| \gg K\} \langle m+n \rangle^{-2}) \right\|_{\ell_{m,n}^2}.
\end{aligned} \tag{2.113}$$

The $\ell_{m,n}^2$ -factor can be estimated by

$$\begin{aligned} & \left\| 1\{|m| \sim M\} 1\{|n| \sim N\} (1\{|m+n| \sim K\} + 1\{|m+n| \gg K\}) \langle m+n \rangle^{-2} \right\|_{\ell_{m,n}^2} \\ & \lesssim M^{\frac{1}{2}} K^{\frac{1}{2}} + M^{\frac{1}{2}} \lesssim M^{\frac{1}{2}} K^{\frac{1}{2}}. \end{aligned}$$

Thus, (2.113) can be estimated by

$$p \sum_{M \sim N \gg 1} \sum_{L, L' \sim M} \sum_{K \geq 1} M^{\alpha - \frac{1}{2}} K^{\alpha - \frac{1}{2}} K^{\frac{1}{q}} \lesssim p.$$

This finishes the proof of the uniform bounds (2.99).

Finally, the convergence statement (2.100) can be deduced by a variant of the preceding arguments. Specifically, we can exploit that in the difference on the left-hand side of (2.100) at least one input must be at frequencies $\geq \varepsilon^{-1}$. This allows us to gain a small power of ε at every step of the preceding argument and therefore yields convergence as $\varepsilon \rightarrow 0$. \square

Remark 2.31. It may be possible to prove Proposition 2.30 by working directly in physical space and with the covariance of the Euclidean Brownian motion W . However, the argument using random Fourier series is closer to the literature on random dispersive equations.

The next proposition establishes the main properties of the smooth approximations of the Brownian paths. As an auxiliary tool we need to introduce a canonical smooth and compactly supported extension $P_{\text{ext}}: \mathbb{R}^D \rightarrow L(\mathbb{R}^D; \mathbb{R}^D)$ of the orthogonal projection P introduced above, i.e., $P_{\text{ext}}(p)$ is the orthogonal projection to $T_p \mathcal{M}$ for all $p \in \mathcal{M}$.

Proposition 2.32 (Smooth approximation of the Brownian paths). *Let $x_0 \in \mathbb{R}$ be arbitrary and let $s < \alpha < \frac{1}{2}$ with $0 < \frac{1}{2} - \alpha \ll 1$ sufficiently small. There exists a small constant $0 < c(\mathcal{M}) \ll 1$ depending only on \mathcal{M} such that for all $0 < \tau \leq c(\mathcal{M})$, there exists an event $\mathcal{E}(\tau, x_0) \subseteq \Omega$ satisfying*

$$\mathbb{P}(\mathcal{E}(\tau, x_0)) \geq 1 - C \exp(-c\tau^{-c})$$

and on which the following properties hold:

- (i) *Convergence:* There exists a path $B_{\tau, x_0, \text{loc}}: \mathbb{R} \rightarrow \mathcal{M}$ such that $B_{\tau, x_0, \text{loc}}^\varepsilon \rightarrow B_{\tau, x_0, \text{loc}}$ in C^α .
- (ii) *Effect of the spatial cut-off:* For all $\varepsilon > 0$ and $x \in [-2, 2]$, it holds that $B_{\tau, x_0, \text{loc}}^\varepsilon(x) = B_{\tau, x_0}^\varepsilon(x)$ and $B_{\tau, x_0, \text{loc}}(x) = B(\tau x + x_0)$.
- (iii) *Para-controlled structure:* For all $\varepsilon > 0$, there exists a vector-valued function $B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} \in C^{2\alpha}$ such that

$$B_{\tau, x_0, \text{loc}}^\varepsilon = \tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) + B_{\tau, x_0, \text{loc}}^{\varepsilon, \#}. \quad (2.114)$$

- (iv) *Uniform bounds:* For all $\varepsilon > 0$, it holds that

$$\|B_{\tau, x_0, \text{loc}}^\varepsilon - B_{\tau, x_0, \text{loc}}^\varepsilon(0)\|_{C^\alpha} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c} \quad \text{and} \quad \|B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0)\|_{C^{2\alpha}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}. \quad (2.115)$$

(v) *Uniform high \times high \rightarrow low-bounds:* For all $1 \leq j_1, j_2 \leq D$, it holds that

$$\sup_{\varepsilon > 0} \sup_{M \sim N} M^s \|P_M(B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1}(0)) \partial_x P_N(B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2}(0))\|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{1-c}. \quad (2.116)$$

(vi) *Convergence of high \times high \rightarrow low-term:* For all $1 \leq j_1, j_2 \leq D$, it holds that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \sup_{M \sim N} M^s \|P_M(B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1}(0)) \partial_x P_N(B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2}(0)) \\ - P_M(B_{\tau, x_0, \text{loc}}^{j_1} - B_{\tau, x_0, \text{loc}}^{j_1}(0)) \partial_x P_N(B_{\tau, x_0, \text{loc}}^{j_2} - B_{\tau, x_0, \text{loc}}^{j_2}(0))\|_{C^{s-1}} = 0. \end{aligned} \quad (2.117)$$

Proof. We begin by recording a Wong-Zakai type approximation result for the Brownian motions $B(x)$, $x \in \mathbb{R}$, on \mathcal{M} defined as solutions to the Stratonovich differential equation (2.93). It follows from [IW89, Chapter 6, Theorem 7.2] that the solutions $B^\varepsilon(x)$, $x \in \mathbb{R}$, to the classical ordinary differential equation (2.94) approximate the Brownian paths $B(x)$, $x \in \mathbb{R}$, in the sense that for any $r > 0$ we have

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\sup_{|x| \leq r} |B^\varepsilon(x) - B(x)|^2 \right] = 0. \quad (2.118)$$

This identifies the Brownian path B as the limit of the approximating sequence $(B^\varepsilon)_{\varepsilon > 0}$ in a weak sense.

We now describe how to obtain an event $\mathcal{E}(\tau, x_0) \subset \Omega$ on which the properties (i)–(vi) hold. To this end we use the para-controlled framework from [GIP15, Theorem 3.3] for the analysis of rough differential equations of the form

$$\partial_x u = F(u)\xi, \quad u(0) = u_0, \quad (2.119)$$

where $u_0 \in \mathbb{R}^d$, $u: \mathbb{R} \rightarrow \mathbb{R}^d$ is a continuous vector-valued function, $\xi: \mathbb{R} \rightarrow \mathbb{R}^n$ is a vector-valued distribution with values in C^β for some $\beta \in (1/3, 1)$, and $F: \mathbb{R}^d \rightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^d)$ is a family of C_b^β vector fields on \mathbb{R}^d . Correspondingly, we view the family of solutions $B_{\tau, x_0, \text{loc}}^\varepsilon: \mathbb{R} \rightarrow \mathcal{M}$, $\varepsilon > 0$, to the ordinary differential equations (2.96) as \mathbb{R}^D -valued solutions to the following rough differential equations

$$\partial_x B_{\tau, x_0, \text{loc}}^\varepsilon = \tau^{\frac{1}{2}} P_{ext}(B_{\tau, x_0, \text{loc}}^\varepsilon) \partial_x (\chi W_{\tau, x_0}^\varepsilon), \quad B_{\tau, x_0, \text{loc}}^\varepsilon(0) = B^\varepsilon(x_0), \quad (2.120)$$

where $P_{ext}: \mathbb{R}^D \rightarrow L(\mathbb{R}^D; \mathbb{R}^D)$ is the canonical smooth and compactly supported extension of the orthogonal projection P introduced above. Clearly, $P_{ext} \in C_b^3$. Observe that while we only consider the initial condition $B_{\tau, x_0, \text{loc}}^\varepsilon(0) = B^\varepsilon(x_0) \in \mathcal{M}$, (2.120) is well-defined for any initial condition in \mathbb{R}^D and falls into the category of rough differential equations (2.119) considered in [GIP15, Theorem 3.3].

In particular, by Lemma 2.30 (with the choice $\lambda = \tau^{-c}$) for any $0 < \frac{1}{2} - \alpha \ll 1$, there exists an event $\mathcal{E}(\tau, x_0) \subset \Omega$ with $\mathbb{P}(\mathcal{E}(\tau, x_0)) \geq 1 - c^{-1} \exp(-c\tau^{-c})$ on which the high-high paraproducts $(\chi(\cdot) W_{\tau, x_0}^\varepsilon) \odot \partial_x (\chi(\cdot) W_{\tau, x_0}^\varepsilon)$ converge⁸ in $C^{2\alpha-1}$ as $\varepsilon \rightarrow 0$, and on which we have for $1 \leq j_1, j_2 \leq D$,

$$\sup_{\varepsilon > 0} \sup_{M \sim N} M^\alpha \|P_M(\chi(\cdot) W_{\tau, x_0}^{\varepsilon, j_1}) \partial_x P_N(\chi(\cdot) W_{\tau, x_0}^{\varepsilon, j_2})\|_{C^{\alpha-1}} \leq \tau^{-c}. \quad (2.121)$$

⁸ This conclusion can be obtained from a slight variant of the proof of Lemma 2.30. We emphasize that the convergence (2.100) in the statement of Lemma 2.30 is stronger than what is needed to conclude the convergence of the high-high paraproducts in $C^{2\alpha-1}$.

Additionally, by a much simpler variant of the proof of Lemma 2.30 we can achieve that on the event $\mathcal{E}(\tau, x_0)$, we have for all integers $1 \leq j \leq D$,

$$\sup_{\varepsilon > 0} \|\chi(\cdot) W_{\tau, x_0}^{\varepsilon, j}\|_{C^\alpha} + \sup_{\varepsilon > 0} \|\partial_x(\chi(\cdot) W_{\tau, x_0}^{\varepsilon, j})\|_{C^{\alpha-1}} \leq \tau^{-\frac{\varepsilon}{2}}. \quad (2.122)$$

Then all assumptions of [GIP15, Theorem 3.3] are satisfied and it follows that there exists $0 < c(\mathcal{M}) \ll 1$ such that for all $0 < \tau \leq c(\mathcal{M})$, we obtain on the event $\mathcal{E}(\tau, x_0)$ the existence of a limit path $B_{\tau, x_0, \text{loc}}: \mathbb{R} \rightarrow \mathcal{M}$ with $B_{\tau, x_0, \text{loc}} \in C^\alpha$ such that $B_{\tau, x_0, \text{loc}}^\varepsilon \rightarrow B_{\tau, x_0, \text{loc}}$ in C^α . By uniqueness of solutions to classical ordinary differential equations, we have $B_{\tau, x_0, \text{loc}}^\varepsilon(x) = B_{\tau, x_0}^\varepsilon(x)$ for all $x \in [-2, 2]$. Additionally, in view of (2.118) we may conclude that $B_{\tau, x_0, \text{loc}}(x) = B(\tau x + x_0)$ for all $x \in [-2, 2]$.

Moreover, it follows⁹ from the proof of [GIP15, Theorem 3.3] that for all $\varepsilon > 0$, there exists a vector-valued function $B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} \in C^{2\alpha}$ such that on $\mathcal{E}(\tau, x_0)$ it holds that

$$B_{\tau, x_0, \text{loc}}^\varepsilon = \tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) + B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} \quad (2.123)$$

with

$$\|B_{\tau, x_0, \text{loc}}^\varepsilon - B_{\tau, x_0, \text{loc}}^\varepsilon(0)\|_{C^\alpha} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c} \quad \text{and} \quad \|B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0)\|_{C^{2\alpha}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}. \quad (2.124)$$

Note that since the target manifold \mathcal{M} is compact, we have $|B_{\tau, x_0, \text{loc}}^\varepsilon(x)| \leq C(\mathcal{M})$ for all $x \in \mathbb{R}$ for some constant $C(\mathcal{M}) \geq 1$, whence the preceding bounds also imply

$$\|B_{\tau, x_0, \text{loc}}^\varepsilon\|_{C^\alpha} \lesssim C(\mathcal{M}). \quad (2.125)$$

Finally, we turn to the proofs of the uniform high \times high \rightarrow low-bounds (2.116) and the convergence of the high \times high \rightarrow low-term (2.117). We focus on the proof of (2.116). The derivation of (2.117) is based on a variant of the argument and is left to the reader. First, we record that by Bony's parilinearization in Besov spaces, see for instance [BCD11, Theorem 2.89 and 2.92], we have the bound

$$\|P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon)\|_{C^\alpha} \leq C(P_{\text{ext}}, \alpha)(1 + \|B_{\tau, x_0, \text{loc}}^\varepsilon\|_{C^\alpha}^{10}) \|B_{\tau, x_0, \text{loc}}^\varepsilon\|_{C^\alpha}, \quad (2.126)$$

where $C(P_{\text{ext}}, \alpha) > 0$ is a constant that depends on α and finitely many derivatives of P_{ext} . In what follows we suppress the superscripts $1 \leq j_1, j_2 \leq D$. Inserting the para-controlled structure (2.123), we find

$$\begin{aligned} & \sup_{\varepsilon > 0} \sup_{M \sim N} M^s \|P_M(B_{\tau, x_0, \text{loc}}^\varepsilon - B_{\tau, x_0, \text{loc}}^\varepsilon(0)) \partial_x P_N(B_{\tau, x_0, \text{loc}}^\varepsilon - B_{\tau, x_0, \text{loc}}^\varepsilon(0))\|_{C^{s-1}} \\ & \leq \sup_{\varepsilon > 0} \sup_{M \sim N} M^s \left\| P_M \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \right) \partial_x P_N \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \right) \right\|_{C^{s-1}} \\ & \quad + \sup_{\varepsilon > 0} \sup_{M \sim N} M^s \left\| P_M \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \right) \partial_x P_N (B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0)) \right\|_{C^{s-1}} \\ & \quad + \sup_{\varepsilon > 0} \sup_{M \sim N} M^s \left\| P_M (B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0)) \partial_x P_N \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon) \otimes (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \right) \right\|_{C^{s-1}} \\ & \quad + \sup_{\varepsilon > 0} \sup_{M \sim N} M^s \|P_M(B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0)) \partial_x P_N(B_{\tau, x_0, \text{loc}}^{\varepsilon, \#} - B_{\tau, x_0, \text{loc}}^\varepsilon(0))\|_{C^{s-1}}. \end{aligned} \quad (2.127)$$

⁹ Specifically, in the notation of the proof of [GIP15, Theorem 3.3] we have $C_F \simeq \tau^{\frac{1}{2}} (\|P_{\text{ext}}\|_{C_b^2}^2 + \|P_{\text{ext}}\|_{C_b^2}^2)$ and $C_\xi \lesssim \tau^{-c}$. To conclude the bounds (2.115) we have to impose $C_F C_\xi \ll 1$.

Then the first term on the right-hand side of (2.127) is the main contribution and all other terms on the right-hand side are easier to estimate due to the better $C^{2\alpha}$ regularity of $B_{\tau,x_0,\text{loc}}^{\varepsilon,\#}$. For this reason we omit the details of their treatment. We also only consider the case when the derivative falls onto the high frequency input in the first term on the right-hand side of (2.127). Recalling that $s < \alpha < \frac{1}{2}$, we decompose this term as follows

$$\begin{aligned}
& \sup_{\varepsilon>0} \sup_{M \sim N} M^s \left\| P_M \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon}) \otimes (\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) P_N \left(\tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon}) \otimes \partial_x (\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) \right\|_{C^{s-1}} \\
& \lesssim \sup_{\varepsilon>0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} M^s \tau \left\| P_M \left(P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) \right. \\
& \quad \times P_N \left(P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \partial_x P_{N'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) \left. \right\|_{C^{\alpha-1}} \\
& \lesssim \sup_{\varepsilon>0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} M^s \tau \left\| P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_M P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right. \\
& \quad \times P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \partial_x P_N P_{N'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \left. \right\|_{C^{\alpha-1}} \\
& \quad + \sup_{\varepsilon>0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} M^s \tau \\
& \quad \times \left\| \left(P_M \left(P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) - P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_M P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) \right. \\
& \quad \times P_N \left(P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \partial_x P_{N'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right) \left. \right\|_{C^{\alpha-1}} \\
& \quad + \{\text{similar}\},
\end{aligned} \tag{2.128}$$

where the quantity {similar} is short-hand notation for two further terms like the second term on the right-hand side of (2.128) with at least one input featuring a commutator term with improved regularity. Then the estimate for the second term on the right-hand side of (2.128) is straightforward using an analogue of the commutator estimate (2.43) for the standard Besov spaces, see for instance [GIP15, Lemma 2.3]. We omit the details. In order to estimate the first term on the right-hand side of (2.128) we pick $0 < \nu \ll 1$ to be fixed sufficiently small further below. Using the standard para-product estimates in Besov spaces, see for instance [GIP15, Lemma 2.1], we find

$$\begin{aligned}
& \sup_{\varepsilon>0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} M^s \tau \left\| P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_M P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right. \\
& \quad \times P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \partial_x P_N P_{N'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \left. \right\|_{C^{\alpha-1}} \\
& \lesssim \sup_{\varepsilon>0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} M^{s-\alpha} \left\| P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \right\|_{C^{1-\alpha+\nu}} \\
& \quad \times M^{\alpha} \tau \left\| P_M P_{M'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \partial_x P_N P_{N'}(\chi(\cdot) W_{\tau,x_0}^{\varepsilon}) \right\|_{C^{\alpha-1}}.
\end{aligned} \tag{2.129}$$

For the $C^{1-\alpha+\nu}$ norm of the product of the low frequency pieces we have

$$\begin{aligned}
& \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} \left\| P_L(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) P_{L'}(P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon})) \right\|_{C^{1-\alpha+\nu}} \\
& \lesssim \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} \sum_{L \ll M'} \sum_{L' \ll N'} \sum_{K \lesssim \max\{L, L'\}} K^{1-\alpha+\nu} L^{-\alpha} (L')^{-\alpha} \left\| P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon}) \right\|_{C^{\alpha}}^2 \\
& \lesssim M^{1-2\alpha+\nu} \left\| P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^{\varepsilon}) \right\|_{C^{\alpha}}^2.
\end{aligned}$$

Note that the estimate (2.121) extends to the setting where each input carries an additional Littlewood-Paley projection to a comparable dyadic frequency region. Thus, using (2.121) and (2.126) we can bound the right-hand side of (2.129) by

$$\begin{aligned}
& \sup_{\varepsilon > 0} \sup_{M \sim N} \sup_{M' \sim M} \sup_{N' \sim N} M^{s+1-3\alpha+\nu} \|P_{\text{ext}}(B_{\tau, x_0, \text{loc}}^\varepsilon)\|_{C^\alpha}^2 \\
& \quad \times M^\alpha \tau \left\| P_M P_{M'} (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \partial_x P_N P_{N'} (\chi(\cdot) W_{\tau, x_0}^\varepsilon) \right\|_{C^{\alpha-1}} \\
& \lesssim C(P_{\text{ext}}, \alpha)^2 (1 + \|B_{\tau, x_0, \text{loc}}^\varepsilon\|_{C^\alpha}^{10})^2 \|B_{\tau, x_0, \text{loc}}^\varepsilon\|_{C^\alpha}^2 \tau^{1-c} \\
& \lesssim_{\mathcal{M}} \tau^{1-c},
\end{aligned}$$

which yields (2.116). Here we used that for any $0 < \frac{1}{2} - s \ll 1$, we may choose $s < \alpha < \frac{1}{2}$ sufficiently close to $\frac{1}{2}$ and we may pick $0 < \nu \ll 1$ sufficiently small so that $s + 1 - 3\alpha + \nu < 0$. This finishes the proof of the proposition. \square

So far, we have constructed the Brownian path $B: \mathbb{R} \rightarrow \mathcal{M}$ and its smooth approximations. We now turn to initial velocities. To this end, we let $\bar{W}: \mathbb{R} \rightarrow \mathbb{R}^D$ be an independent copy of the Euclidean Brownian motion W . Then, we define the integrated velocity $\mathcal{V}: \mathbb{R} \rightarrow \mathbb{R}^D$ by the Stratonovich integral¹⁰

$$\mathcal{V}(x) := \int_0^x P(B(y)) \circ d\bar{W}(y). \quad (2.130)$$

We define the velocity $V: \mathbb{R} \rightarrow \mathbb{R}^D$ as the distributional derivative of \mathcal{V} . We now define smooth approximations of both \mathcal{V} and V . To this end, we first define a smooth approximation of the Euclidean Brownian motion by

$$\bar{W}^\varepsilon := P_{\leq \varepsilon^{-1}} \bar{W} = K_\varepsilon * \bar{W}.$$

Then, we define the velocity $V^\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^D$ and its integral \mathcal{V}^ε by

$$V^\varepsilon(x) := P(B^\varepsilon(x)) \partial_x \bar{W}^\varepsilon(x), \quad (2.131)$$

$$\mathcal{V}^\varepsilon(x) := \int_0^x V^\varepsilon(y) dy. \quad (2.132)$$

We note that $V^\varepsilon \in (B^\varepsilon)^* T\mathcal{M}$, i.e., $V^\varepsilon(x) \in T_{B^\varepsilon(x)} \mathcal{M}$ for all $x \in \mathbb{R}$. Next, for any $\tau > 0$ and $x_0 \in \mathbb{R}$ we define the re-scaled and translated velocity, respectively integrated velocity, by

$$V_{\tau, x_0}^\varepsilon(x) := \tau V^\varepsilon(\tau x + x_0), \quad \mathcal{V}_{\tau, x_0}^\varepsilon(x) := \mathcal{V}^\varepsilon(\tau x + x_0) - \mathcal{V}^\varepsilon(x_0). \quad (2.133)$$

Introducing the associated re-scaled and translated Euclidean Brownian motion

$$\bar{W}_{\tau, x_0}^\varepsilon(x) := \tau^{-\frac{1}{2}} (\bar{W}^\varepsilon(\tau x + x_0) - \bar{W}^\varepsilon(x_0)),$$

it follows that $V_{\tau, x_0}^\varepsilon$ satisfies

$$V_{\tau, x_0}^\varepsilon(x) = \tau^{\frac{1}{2}} P(B_{\tau, x_0}^\varepsilon(x)) \partial_x \bar{W}_{\tau, x_0}^\varepsilon(x).$$

¹⁰ Since B and \bar{W} are independent and hence have vanishing cross-variation, the Itô and Stratonovich integrals are identical in this case.

Furthermore, we define the spatially localized versions $\mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon$ and $V_{\tau,x_0,\text{loc}}^\varepsilon$ by

$$V_{\tau,x_0,\text{loc}}^\varepsilon(x) := \tau^{\frac{1}{2}} P(B_{\tau,x_0,\text{loc}}^\varepsilon(x)) \partial_x (\chi(x) \bar{W}_{\tau,x_0}^\varepsilon(x)), \quad (2.134)$$

$$\mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon(x) := \int_0^x V_{\tau,x_0,\text{loc}}^\varepsilon(y) dy, \quad (2.135)$$

where the smooth cutoff $\chi(x)$ is the same as above. Note that $V_{\tau,x_0}^\varepsilon \in (B_{\tau,x_0}^\varepsilon)^* T\mathcal{M}$ and $V_{\tau,x_0,\text{loc}}^\varepsilon \in (B_{\tau,x_0,\text{loc}}^\varepsilon)^* T\mathcal{M}$. In particular, we obtain that $V_{\tau,x_0,\text{loc}}^\varepsilon(x) \in T_{B_{\tau,x_0}^\varepsilon(x)}\mathcal{M}$ for $x \in [-2, 2]$.

The main properties of the smooth approximations of the velocity and the integrated velocity are included in the following proposition.

Proposition 2.33. (Smooth approximation of the white noise velocities) *Let $x_0 \in \mathbb{R}$ be arbitrary and let $s < \alpha < \frac{1}{2}$ with $0 < \frac{1}{2} - \alpha \ll 1$ sufficiently small. There exists a small constant $0 < c(\mathcal{M}) \ll 1$ depending only on \mathcal{M} such that for all $0 < \tau \leq c(\mathcal{M})$, there exists an event $\mathcal{E}(\tau, x_0) \subseteq \Omega$ satisfying*

$$\mathbb{P}(\mathcal{E}(\tau, x_0)) \geq 1 - C \exp(-c\tau^{-c})$$

and on which the following properties hold:

- (i) *Convergence: There exist $V_{\tau,x_0,\text{loc}}, \mathcal{V}_{\tau,x_0,\text{loc}}: \mathbb{R} \rightarrow \mathbb{R}^D$ such that $\partial_x \mathcal{V}_{\tau,x_0,\text{loc}} = V_{\tau,x_0,\text{loc}}$ and $(V_{\tau,x_0,\text{loc}}^\varepsilon, \mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon) \rightarrow (V_{\tau,x_0,\text{loc}}, \mathcal{V}_{\tau,x_0,\text{loc}})$ in $C^{\alpha-1} \times C^\alpha$.*
(ii) *Effect of spatial cut-off: For all $\varepsilon > 0$ and $x \in [-2, 2]$, we have*

$$V_{\tau,x_0,\text{loc}}^\varepsilon(x) = V_{\tau,x_0}^\varepsilon(x), \quad V_{\tau,x_0,\text{loc}}(x) = \tau V(\tau x + x_0), \quad (2.136)$$

$$\mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon(x) = \mathcal{V}_{\tau,x_0}^\varepsilon(x), \quad \mathcal{V}_{\tau,x_0,\text{loc}}(x) = \mathcal{V}(\tau x + x_0) - \mathcal{V}(x_0), \quad (2.137)$$

where (2.136) holds in the sense of distributions.

- (iii) *Para-controlled structure: For all $\varepsilon > 0$ there exist vector-valued functions $V_{\tau,x_0,\text{loc}}^{\varepsilon,\#} \in C^{2\alpha-1}$ and $\mathcal{V}_{\tau,x_0,\text{loc}}^{\varepsilon,\#} \in C^{2\alpha}$ such that*

$$\begin{aligned} V_{\tau,x_0,\text{loc}}^\varepsilon &= \tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^\varepsilon) \otimes \partial_x (\chi(\cdot) \bar{W}_{\tau,x_0,\text{loc}}^\varepsilon) + V_{\tau,x_0,\text{loc}}^{\varepsilon,\#}, \\ \mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon &= \tau^{\frac{1}{2}} P_{\text{ext}}(B_{\tau,x_0,\text{loc}}^\varepsilon) \otimes (\chi(\cdot) \bar{W}_{\tau,x_0,\text{loc}}^\varepsilon) + \mathcal{V}_{\tau,x_0,\text{loc}}^{\varepsilon,\#}. \end{aligned} \quad (2.138)$$

- (iv) *Uniform bounds: For all $\varepsilon > 0$, it holds that*

$$\|V_{\tau,x_0,\text{loc}}^\varepsilon\|_{C^{\alpha-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \quad \|\mathcal{V}_{\tau,x_0,\text{loc}}^\varepsilon\|_{C^\alpha} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \quad (2.139)$$

$$\|V_{\tau,x_0,\text{loc}}^{\varepsilon,\#}\|_{C^{2\alpha-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \quad \|\mathcal{V}_{\tau,x_0,\text{loc}}^{\varepsilon,\#}\|_{C^{2\alpha}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}. \quad (2.140)$$

- (v) *Uniform high \times high \rightarrow low-bounds: For all $1 \leq j_1, j_2 \leq D$ and all $\varepsilon > 0$, it holds that*

$$\begin{aligned} \sup_{\varepsilon > 0} \sup_{M \sim N} M^s &\|P_M \mathcal{V}_{\tau,x_0,\text{loc}}^{\varepsilon,j_1} P_N V_{\tau,x_0,\text{loc}}^{\varepsilon,j_2}\|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \\ \sup_{\varepsilon > 0} \sup_{M \sim N} M^s &\|P_M (B_{\tau,x_0,\text{loc}}^{\varepsilon,j_1} - B_{\tau,x_0,\text{loc}}^{\varepsilon,j_1}(0)) P_N V_{\tau,x_0,\text{loc}}^{\varepsilon,j_2}\|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \\ \sup_{\varepsilon > 0} \sup_{M \sim N} M^s &\|P_M \mathcal{V}_{\tau,x_0,\text{loc}}^{\varepsilon,j_1} \partial_x P_N (B_{\tau,x_0,\text{loc}}^{\varepsilon,j_2} - B_{\tau,x_0,\text{loc}}^{\varepsilon,j_2}(0))\|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}. \end{aligned} \quad (2.141)$$

(vi) *Convergence of high \times high \rightarrow low-term: For all $1 \leq j_1, j_2 \leq D$ and all $\varepsilon > 0$, it holds that*

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \sup_{M \sim N} M^s & \| P_M \mathcal{V}_{\tau, x_0, \text{loc}}^{\varepsilon, j_1} P_N V_{\tau, x_0, \text{loc}}^{\varepsilon, j_2} - P_M \mathcal{V}_{\tau, x_0, \text{loc}}^{j_1} P_N V_{\tau, x_0, \text{loc}}^{j_2} \|_{C^{s-1}} = 0, \\ \lim_{\varepsilon \rightarrow 0} \sup_{M \sim N} M^s & \| P_M (B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_1}(0)) P_N V_{\tau, x_0, \text{loc}}^{\varepsilon, j_2} \\ & - P_M (B_{\tau, x_0, \text{loc}}^{j_1} - B_{\tau, x_0, \text{loc}}^{j_1}(0)) P_N V_{\tau, x_0, \text{loc}}^{j_2} \|_{C^{s-1}} = 0, \quad (2.142) \\ \lim_{\varepsilon \rightarrow 0} \sup_{M \sim N} M^s & \| P_M \mathcal{V}_{\tau, x_0, \text{loc}}^{\varepsilon, j_1} P_N (B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2} - B_{\tau, x_0, \text{loc}}^{\varepsilon, j_2}(0)) \\ & - P_M \mathcal{V}_{\tau, x_0, \text{loc}}^{j_1} P_N (B_{\tau, x_0, \text{loc}}^{j_2} - B_{\tau, x_0, \text{loc}}^{j_2}(0)) \|_{C^{s-1}} = 0. \end{aligned}$$

Proof. The assertions can be proved using similar arguments as in the proof of the preceding Proposition 2.32. We refer to [IW89, Chapter 6, Theorem 7.1] for a Wong-Zakai type result for approximations of Stratonovich stochastic integrals. Most of the arguments are in fact simpler because the integrated velocities are defined via stochastic integration, while the Brownian paths are defined in terms of a stochastic differential equation. The details are left to the reader. \square

Corollary 2.34. *Let $B: \mathbb{R} \rightarrow \mathcal{M}$ be the Brownian path, let $V \in B^*T\mathcal{M}$ be the white noise velocity, and let $(B^\varepsilon)_{\varepsilon>0}$ and $(V^\varepsilon)_{\varepsilon>0}$ be their smooth approximations. For any $R > 0$ we have almost surely that $(B^\varepsilon, V^\varepsilon)$ converges to (B, V) in $C^s \times C^{s-1}([-R, R]) \rightarrow T\mathcal{M}$ as $\varepsilon \rightarrow 0$.*

Proof. The assertion follows by a covering argument from the properties (i)-(ii) of the smooth local approximations established in Proposition 2.32 and in Proposition 2.33. \square

We recall that the (shifted) linear waves corresponding to the initial data $(B_{\tau, x_0, \text{loc}}^\varepsilon, V_{\tau, x_0, \text{loc}}^\varepsilon)$ are given by

$$\phi_{\tau, x_0, \text{loc}}^{\pm, \varepsilon} := \frac{1}{2} \left(B_{\tau, x_0, \text{loc}}^\varepsilon - B_{\tau, x_0, \text{loc}}^\varepsilon(0) \mp \mathcal{V}_{\tau, x_0, \text{loc}}^\varepsilon \right),$$

with analogous definitions of the linear waves $\phi_{\tau, x_0, \text{loc}}^\pm$ for the initial data $(B_{\tau, x_0, \text{loc}}, V_{\tau, x_0, \text{loc}})$. In the final proposition of this section, we prove a variant of the high \times high \rightarrow low-bound, which also allows for shifts in the argument.

Proposition 2.35. (High \times high \rightarrow low-bound with shifts) *Let $x_0 \in \mathbb{R}$ be arbitrary and let $s < \alpha < \frac{1}{2}$ with $0 < \frac{1}{2} - \alpha \ll 1$ sufficiently small. There exists a small constant $0 < c(\mathcal{M}) \ll 1$ depending only on \mathcal{M} such that for all $0 < \tau \leq c(\mathcal{M})$, there exists an event $\mathcal{E}(\tau, x_0) \subseteq \Omega$ satisfying*

$$\mathbb{P}(\mathcal{E}(\tau, x_0)) \geq 1 - C \exp(-c\tau^{-c})$$

and on which the following properties hold:

(i) *Uniform high \times high \rightarrow low-bounds: For all $1 \leq j_1, j_2 \leq D$, it holds that*

$$\begin{aligned} \sup_{\varepsilon > 0} \sup_{t \in \mathbb{R}} \sup_{M \sim N} M^s & \| P_M \phi_{\tau, x_0, \text{loc}}^{+, \varepsilon, j_1}(x-t) \partial_x P_N \phi_{\tau, x_0, \text{loc}}^{-, \varepsilon, j_2}(x+t) \|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}, \\ \sup_{\varepsilon > 0} \sup_{t \in \mathbb{R}} \sup_{M \sim N} M^s & \| \partial_x P_M \phi_{\tau, x_0, \text{loc}}^{+, \varepsilon, j_1}(x-t) P_N \phi_{\tau, x_0, \text{loc}}^{-, \varepsilon, j_2}(x+t) \|_{C^{s-1}} \lesssim_{\mathcal{M}} \tau^{\frac{1}{2}-c}. \end{aligned} \quad (2.143)$$

(ii) *Convergence of high \times high \rightarrow low-term:* For all $1 \leq j_1, j_2 \leq D$, it holds that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \sup_{t \in \mathbb{R}} \sup_{M \sim N} M^s \left\| P_M \phi_{\tau, x_0, \text{loc}}^{+, \varepsilon, j_1}(x-t) \partial_x P_N \phi_{\tau, x_0, \text{loc}}^{-, \varepsilon, j_2}(x+t) \right. \\ \left. - P_M \phi_{\tau, x_0, \text{loc}}^{+, j_1}(x-t) \partial_x P_N \phi_{\tau, x_0, \text{loc}}^{-, j_2}(x+t) \right\|_{C^{s-1}} = 0. \\ \lim_{\varepsilon \rightarrow 0} \sup_{t \in \mathbb{R}} \sup_{M \sim N} M^s \left\| \partial_x P_M \phi_{\tau, x_0, \text{loc}}^{+, \varepsilon, j_1}(x-t) P_N \phi_{\tau, x_0, \text{loc}}^{-, \varepsilon, j_2}(x+t) \right. \\ \left. - \partial_x P_M \phi_{\tau, x_0, \text{loc}}^{+, j_1}(x-t) P_N \phi_{\tau, x_0, \text{loc}}^{-, j_2}(x+t) \right\|_{C^{s-1}} = 0. \end{aligned} \quad (2.144)$$

Remark 2.36. In Proposition 2.35, it is essential to consider the mixed term involving both ϕ^+ and ϕ^- . In other cases, the probabilistic resonance poses a severe problem.

Proof. The argument is similar to the proofs of Proposition 2.32 and Proposition 2.33, but this time we get no resonant term (at top order). The reason is that $W + \bar{W}$ and $W - \bar{W}$ are independent due to the rotation invariance of Brownian motion. To see this more directly, we check that the cross-covariance matrix vanishes, which suffices to conclude independence for Gaussian random variables. Indeed, we have for any $x_1, x_2 \in \mathbb{R}$ and any $1 \leq i, j \leq D$ that

$$\begin{aligned} & \text{cov}((W^i + \bar{W}^i)(x_1), (W^j - \bar{W}^j)(x_2)) \\ &= \mathbb{E} \left[\left((W^i + \bar{W}^i)(x_1) - \mathbb{E}[(W^i + \bar{W}^i)(x_1)] \right) \left((W^j - \bar{W}^j)(x_2) - \mathbb{E}[(W^j - \bar{W}^j)(x_2)] \right) \right] \\ &= \mathbb{E} \left[\left(W^i(x_1) - \mathbb{E}[W^i(x_1)] \right) \left(W^j(x_2) - \mathbb{E}[W^j(x_2)] \right) \right] \\ &\quad - \mathbb{E} \left[\left(\bar{W}^i(x_1) - \mathbb{E}[\bar{W}^i(x_1)] \right) \left(\bar{W}^j(x_2) - \mathbb{E}[\bar{W}^j(x_2)] \right) \right] \\ &= 0, \end{aligned}$$

as desired. \square

3. Ansatz

In the introduction (Sect. 1.3), we gave an informal description of the Ansatz for ϕ . In this section, we give a rigorous definition of the terms in our previous discussion.

3.1. The shifted wave map. In the proof of Theorem 1.4, which is a local result, we utilize some form of smallness. In principle, smallness can be created by localizing to times $|t| \leq \tau \ll 1$. However, due to the complexity of our Ansatz, exhibiting gains in τ in this fashion is tedious. Instead of localizing to a short time-interval, we use a scaling argument. To this end, we use the scale-invariance, translation-invariance, and finite speed of propagation of (WM), which allows us to replace B^ε and V^ε by their re-scaled, translated, and localized counterparts (see, e.g., the definition of $B_{\tau, x_0, \text{loc}}^\varepsilon$ in (2.96) and the definition of $V_{\tau, x_0, \text{loc}}^\varepsilon$ in (2.134)). Unfortunately, this alone is not sufficient to yield small data, since the manifold constraint $B^\varepsilon(x) \in \mathcal{M}$ prohibits smallness in L^∞ . To resolve this issue, we need to subtract the initial position at $x = 0$, and the resulting shifted wave map will be described in this section.

While our main result (Theorem 1.4) is stated just for the random data $(B^\varepsilon, V^\varepsilon)$, and will be proven using the re-scaled, translated, and localized counterparts $B_{\tau, x_0, \text{loc}}^\varepsilon$ and $V_{\tau, x_0, \text{loc}}^\varepsilon$, our argument requires only the regularity and high \times high \rightarrow low-estimates from Sect. 2.6. As a result, it is more convenient to work in a general framework. To this end, we let $\phi_0: \mathbb{R} \rightarrow \mathcal{M}$ and $\phi_1 \in \phi_0^* T\mathcal{M}$. Then, we define the shifted initial data by

$$\phi_0^\diamond(x) := \phi_0(x) - \phi_0(0) \quad \text{and} \quad \phi_1^\diamond(x) = \phi_1(x). \quad (3.1)$$

We remark that the shifted data satisfies the geometric constraint

$$\phi_1^\diamond \in (\phi_0^\diamond)^* T(\mathcal{M} - \phi_0(0)),$$

but this will not be used in the rest of the paper. We define the corresponding shifted linear waves $\phi^{\diamond,+}$ and $\phi^{\diamond,-}$ by

$$\phi^{\diamond,\pm}(x) = \frac{1}{2\theta} \left(\phi_0^\diamond(x) \mp \int_0^x dy \phi_1^\diamond(y) \right). \quad (3.2)$$

Here, the parameter $\theta > 0$ is as in Sect. 2.1. The purposes of the $(1/\theta)$ -factor in (3.2) will be clear from (3.6) below. We also define the shifted second fundamental form by

$$\mathbb{I}_{ij}^{\diamond,k}(\zeta) = \mathbb{I}_{ij}^{\diamond,k}(\zeta; \phi_0(0)) := \mathbb{I}_{ij}^k(\zeta + \phi_0(0)). \quad (3.3)$$

For most of this paper, the dependence of $\mathbb{I}_{ij}^{\diamond,k}$ on $\phi_0(0)$ is omitted from our notation. Finally, we define the shifted wave map ϕ^\diamond by

$$\phi^\diamond(t, x) := \phi(t, x) - \phi_0(0). \quad (3.4)$$

As a result of our definitions, we see that ϕ^\diamond satisfies the shifted wave maps equation

$$\begin{cases} \partial_\mu \partial^\mu \phi^{\diamond,k} = -\mathbb{I}_{ij}^{\diamond,k}(\phi^\diamond) \partial_\mu \phi^{\diamond,i} \partial^\mu \phi^{\diamond,j} \\ \phi^\diamond(0, x) = \phi_0^\diamond(x), \quad \partial_t \phi^\diamond(0, x) = \phi_1^\diamond(x). \end{cases} \quad (3.5)$$

In our analysis of (3.5), we will exclusively work with its Duhamel integral formulation. Due to the mapping properties of the Duhamel integral on Hölder spaces, it is convenient to also insert our cut-off functions χ^+ and χ^- from (2.8). Due to finite speed of propagation, this does not alter the behavior of ϕ^\diamond on $|u|, |v| \leq 2$. In total, this leads to the fixed-point problem

$$\phi^{\diamond,k} = \theta \phi^{\diamond,+k} + \theta \phi^{\diamond,-k} - \chi^+ \chi^- \text{Duh} \left[\mathbb{I}_{ij}^{\diamond,k}(\phi^\diamond) \partial_u \phi^{\diamond,i} \partial_v \phi^{\diamond,j} \right]. \quad (3.6)$$

As stated above, the purpose of the $(1/\theta)$ -factor in (3.2) stems from (3.6). It leads to the small θ -factor in (3.6), which allows us to consider small modulations in our Ansatz (see Sect. 3.3).

A word on notation: Throughout Sects. 3–8, we work exclusively with the shifted wave map ϕ^\diamond and the shifted linear waves $\phi^{\diamond,+}$ and $\phi^{\diamond,-}$. In order to simplify the notation (see e.g. Section 6), we omit the superscript “ \diamond ” in our notation for ϕ^\diamond , $\phi^{\diamond,+}$, and $\phi^{\diamond,-}$. Therefore, we simply write ϕ , ϕ^+ , and ϕ^- . In Sect. 9, where both the original and shifted wave maps are present, we distinguish between ϕ and ϕ^\diamond . Throughout the whole paper, we include the superscript “ \diamond ” in \mathbb{I} . This is because we do not want to change the definition of a central object in mathematics.

3.2. The enhanced initial data. We let ϕ^+ and ϕ^- be the linear waves from (3.2), where we recall that the superscript “ \diamond ” is omitted from our notation. As will be further detailed in Hypothesis 3.8, our analysis will be based on the following assumptions.

- (1) Regularity condition: It holds that $\|\phi^+(x)\|_{C_x^s(\mathbb{R})} + \|\phi^-(x)\|_{C_x^s(\mathbb{R})} \leq \theta$.
 (2) High \times high \rightarrow low-bounds: For $\pm_1, \pm_2 \in \{+, -\}$ it holds that

$$\begin{aligned} \max_{\pm_1, \pm_2} \max_{1 \leq m, n \leq D} \sup_{M \sim N} M^s \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{C_x^{s-1}(\mathbb{R})} &\leq \theta^2, \\ \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sup_{t \in \mathbb{R}} M^s \|P_M \phi^{+, m}(x-t) \partial_x P_N \phi^{-, n}(x+t)\|_{C_x^{s-1}(\mathbb{R})} &\leq \theta^2, \\ \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sup_{t \in \mathbb{R}} M^s \|\partial_x P_M \phi^{+, m}(x-t) P_N \phi^{-, n}(x+t)\|_{C_x^{s-1}(\mathbb{R})} &\leq \theta^2. \end{aligned}$$

As previously discussed in Remark 2.36, we emphasize that the term involving the supremum over $t \in \mathbb{R}$ only covers products of linear waves in different directions. Due to the bilinear nature of the high \times high \rightarrow low-bound, the two assumptions cannot be captured through a norm or metric on $C_x^s \times C_x^s$. As in the previous literature on random dispersive equations (see e.g. the survey [BOP19]), we will interpret the assumptions through an enhanced data set.

Definition 3.1. (Enhanced data and \mathcal{D}^s) For any $\phi^+, \phi^- : \mathbb{R} \rightarrow \mathbb{R}^D$, $\pm_1, \pm_2 \in \{+, -\}$, $1 \leq m, n \leq D$, and $M, N \geq 1$ satisfying $M \sim N$, we define $\Phi_{M, N}^{\pm_1, \pm_2, m, n} : \mathbb{R} \rightarrow \mathbb{R}$ by

$$\Phi_{M, N}^{\pm_1, \pm_2, m, n}(x) := M^s P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x). \quad (3.7)$$

In the case $\pm_1 \neq \pm_2$, we also define the shifted $\Phi_{M, N}^{\pm_1, \pm_2, (s), m, n} : \mathbb{R}_{t, x}^{1+1} \rightarrow \mathbb{R}$ by

$$\Phi_{M, N}^{\pm_1, \pm_2, (s), m, n}(x; t) := M^s P_M \phi^{\pm_1, m}(x-t) \partial_x P_N \phi^{\pm_2, n}(x+t). \quad (3.8)$$

We then define

$$\begin{aligned} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s} := \max &\left(\|\phi^+\|_{C_x^s} + \|\phi^-\|_{C_x^s}, \max_{\pm_1, \pm_2} \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sqrt{\|\Phi_{M, N}^{\pm_1, \pm_2, m, n}(x)\|_{C_x^{s-1}}}, \right. \\ &\left. \max_{\pm_1 \neq \pm_2} \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sup_{t \in \mathbb{R}} \sqrt{\|\Phi_{M, N}^{\pm_1, \pm_2, (s), m, n}(x; t)\|_{C_x^{s-1}}} \right). \end{aligned}$$

We note that the regularity condition and high \times high \rightarrow low-bound are equivalent to the \mathcal{D}^s -bound $\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta$.

Remark 3.2. As mentioned above, $\|\cdot\|_{\mathcal{D}^s}$ is not actually a norm. However, it can be viewed as the composition of the nonlinear function $\phi^\pm \mapsto (\phi^\pm, \Phi)$, an honest norm, and the nonlinear function $(a, b) \mapsto a + \sqrt{b}$. By including the square-root, we have preserved the 1-homogeneity in ϕ^\pm but destroyed the 1-homogeneity in Φ . While $\|\cdot\|_{\mathcal{D}^s}$ can therefore also not be viewed as a honest norm on the enhanced data (ϕ^\pm, Φ) , we will now associate it with a metric on the enhanced data.

In order to prove the convergence statement in Theorem 1.4, it is convenient to introduce continuous functions on \mathcal{D}^s . To this end, we first define the distance

$$\begin{aligned} &\|(\phi_1^+, \phi_1^-); (\phi_2^+, \phi_2^-)\|_{\mathcal{D}^s} \\ &:= \max \left(\|\phi_1^+ - \phi_2^+\|_{C_x^s} + \|\phi_1^- - \phi_2^-\|_{C_x^s}, \right. \end{aligned}$$

$$\begin{aligned} & \max_{\pm_1, \pm_2} \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sqrt{\|\Phi_{1,M,N}^{\pm_1, \pm_2, m, n}(x) - \Phi_{2,M,N}^{\pm_1, \pm_2, m, n}(x)\|_{C_x^{s-1}}}, \\ & \max_{\pm_1 \neq \pm_2} \max_{1 \leq m, n \leq D} \sup_{M \sim N} \sup_{t \in \mathbb{R}} \sqrt{\|\Phi_{1,M,N}^{\pm_1, \pm_2, (s), m, n}(x; t) - \Phi_{2,M,N}^{\pm_1, \pm_2, (s), m, n}(x; t)\|_{C_x^{s-1}}}. \end{aligned} \quad (3.9)$$

Since square-roots preserve metrics, one can view $\|\cdot\|_{\mathcal{D}^s}$ as a metric on the enhanced data (ϕ^\pm, Φ) .

Definition 3.3 (*Lipschitz continuous dependence w.r.t. \mathcal{D}^s*). Let (\mathbb{X}, d) be a metric space and consider a function $F: C_b^\infty \times C_b^\infty \rightarrow \mathbb{X}$. We say that F is Lipschitz continuous w.r.t. \mathcal{D}^s if there exists a constant $\text{Lip}(F) \in [0, \infty)$ such that

$$d(F(\phi_1^+, \phi_1^-), F(\phi_2^+, \phi_2^-)) \leq \text{Lip}(F) \|(\phi_1^+, \phi_1^-); (\phi_2^+, \phi_2^-)\|_{\mathcal{D}^s} \quad (3.10)$$

for all $(\phi_1^+, \phi_1^-), (\phi_2^+, \phi_2^-) \in \mathcal{D}^s$. In other words, F is Lipschitz continuous as a function of (ϕ^+, ϕ^-, Φ) .

While $\|\cdot\|_{\mathcal{D}^s}$ encapsulates the high \times high \rightarrow low-bound, it does not explicitly contain bounds on the low \times high and high \times low-para-products. The reason is that the optimal bounds for both of them follow directly from the regularity condition $\phi^+, \phi^- \in C_x^s$. For the rest of the paper, it is convenient to encapsulate estimates of all frequency interactions in a single lemma.

Lemma 3.4. *Let $\pm_1, \pm_2 \in \{+, -\}$, let $1 \leq m, n \leq D$, and let M and N be any frequency scales. Then, it holds that*

$$\begin{aligned} & \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{C_x^{r-1}} \\ & \lesssim \left(1\{M \lesssim N\} M^{-s} N^{r-s} + 1\{M \gg N\} M^{r-1-s} N^{1-s}\right) \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2 \\ & \lesssim M^{-s} N^{r-s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2. \end{aligned} \quad (3.11)$$

Proof. The second inequality in (3.11) directly follows from

$$1\{M \gg N\} M^{r-1-s} N^{1-s} \lesssim N^{r-1} M^{-s} N^{1-s} = M^{-s} N^{r-s}.$$

Thus, it remains to prove the first inequality. To this end, we distinguish the two cases $M \sim N$ and $M \not\sim N$. If $M \sim N$, we have that

$$\begin{aligned} \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{C_x^{r-1}} & \lesssim M^{r-s} \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{C_x^{s-1}} \\ & \lesssim M^{r-2s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2 \\ & \lesssim M^{-s} N^{r-s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2. \end{aligned}$$

If $M \not\sim N$, we have that

$$\begin{aligned} \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{C_x^{r-1}} & \lesssim \max(M, N)^{r-1} \|P_M \phi^{\pm_1, m}(x) \partial_x P_N \phi^{\pm_2, n}(x)\|_{L_x^\infty} \\ & \lesssim \max(M, N)^{r-1} \|P_M \phi^{\pm_1, m}(x)\|_{L_x^\infty} \|\partial_x P_N \phi^{\pm_2, n}(x)\|_{L_x^\infty} \\ & \lesssim \max(M, N)^{r-1} M^{-s} N^{1-s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2. \end{aligned}$$

By further distinguishing the cases $M \ll N$ and $M \gg N$, this yields the desired estimate. \square

3.3. Rigorous definition of the Ansatz. In the introduction, we motivated the form of our Ansatz, which will be written as

$$\begin{aligned} \phi^i(u, v) = & \sum_M A_{M,m}^{+,i}(u, v) \phi_M^{+,m}(u) + \sum_N A_{N,n}^{-,i}(u, v) \phi_N^{-,n}(v) \\ & + \sum_{\substack{M,N: \\ M \sim_\delta N}} B_{M,N,mn}^i(u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v) + \psi^i(u, v). \end{aligned} \quad (3.12)$$

Throughout the paper, we impose the following frequency-support conditions on A^+ , A^- , and B .

Condition 3.5 (Frequency-support conditions). *For all $1 \leq i, m, n \leq D$ and frequency-scales M and N , we impose the following frequency-support conditions:*

$$P_{\gg M^{1-\sigma}}^u A_{M,m}^{+,i} = 0, \quad (3.13)$$

$$P_{\gg N^{1-\sigma}}^v A_{N,n}^{-,i} = 0, \quad (3.14)$$

$$P_{\gg M^{1-\sigma}}^u B_{M,N,mn}^i = P_{\gg N^{1-\sigma}}^v B_{M,N,mn}^i = 0. \quad (3.15)$$

While $A_{M,m}^{+,i}$ has u -frequencies $\lesssim M^{1-\sigma}$, we impose no frequency restrictions in the v -variable. This is because high \times high-products such as $\phi_M^{+,m}(u) P_M^v \zeta(v)$, where $\zeta \in C_v^r$, only belong to $C_u^s C_v^r$ but do not belong to $C_u^r C_v^s$.

The equations for A^+ , A^- , B , and ψ will be given further below in Sect. 3.4. For now, we concentrate on the norms in which the four unknowns will be measured. While the modulation equations for A^+ and A^- will eventually be solved using a para-controlled approach (see Sect. 8), the PDE-analysis in Sect. 4–7 only requires the following norms.

Definition 3.6 (Modulation norms). Let $A^+ = (A_{M,m}^{+,i})$, $A^- = (A_{N,n}^{-,j})$, and $B = (B_{M,N,mn}^k)$. Then, we define

$$\|(A^+, A^-, B)\|_{Mod} := \sup_M \|A_M^+\|_{Mod_M^+} + \sup_N \|A_N^-\|_{Mod_N^-} + \max_{1 \leq k, m, n \leq D} \sup_{M \sim_\delta N} \|B_{M,N,mn}^k\|_{C_u^s C_v^s}, \quad (3.16)$$

where

$$\|A_M^+\|_{Mod_M^+} = \max_{1 \leq i, m \leq D} \left(\|P_{\lesssim M^{1-\delta}}^v A_{M,m}^{+,i}\|_{C_u^s C_v^s} + \|P_{\gg M^{1-\delta}}^v A_{M,m}^{+,i}\|_{C_u^s C_v^r} \right), \quad (3.17)$$

$$\|A_N^-\|_{Mod_N^-} = \max_{1 \leq j, n \leq D} \left(\|P_{\lesssim N^{1-\delta}}^u A_{N,n}^{-,j}\|_{C_u^s C_v^s} + \|P_{\gg N^{1-\delta}}^u A_{N,n}^{-,j}\|_{C_u^r C_v^s} \right). \quad (3.18)$$

Using the definition of the modulation norms, we directly obtain the following estimates for A^+ and A^- .

Corollary 3.7 (δ -gain for A^\pm). *For all dyadic scales $M, N \geq 1$, it holds that*

$$\|\partial_v A_M^+\|_{C_u^s C_v^{r-1}} \lesssim M^{(1-\delta)(r-s)} \|A_M^+\|_{Mod_M^+}, \quad (3.19)$$

$$\|\partial_u A_N^-\|_{C_u^{r-1} C_v^s} \lesssim N^{(1-\delta)(r-s)} \|A_N^-\|_{Mod_N^-}. \quad (3.20)$$

We refer to the $(1 - \delta)$ -factor in the exponents as a δ -gain. On a microscopic level, it is a direct consequence of the modified frequency-cutoffs in (3.17) and (3.18). On a macroscopic level, the δ -gain is made possible by including the bilinear term $B_{M,N,mn}^i \phi_M^{+,m} \phi_N^{-,n}$ in our Ansatz.

Proof. It suffices to prove (3.19), since the argument for (3.20) is similar. For all M , we have that

$$\begin{aligned} \|\partial_v A_M^+\|_{C_u^s C_v^{r-1}} &\leq \|\partial_v P_{\lesssim M^{1-\delta}}^v A_M^+\|_{C_u^s C_v^{r-1}} + \|\partial_v P_{\gg M^{1-\delta}}^v A_M^+\|_{C_u^s C_v^{r-1}} \\ &\lesssim M^{(1-\delta)(r-s)} \|P_{\lesssim M^{1-\delta}}^v A_M^+\|_{C_u^s C_v^s} + \|P_{\gg M^{1-\delta}}^v A_M^+\|_{C_u^s C_v^r} \\ &\lesssim M^{(1-\delta)(r-s)} \|A_M^+\|_{Mod_M^+}. \end{aligned}$$

□

In Sect. 4.8, we will prove numerous estimates for the components in our Ansatz. To this end, it is helpful to collect all assumptions on the initial data and components in a single hypothesis, which can then be referenced easily. This hypothesis will ultimately form the basis of a contraction mapping argument.

Hypothesis 3.8 (Smallness). *Let θ be the smallness parameter introduced in Sect. 2.1. Then, we assume that*

$$\max \left(\|(\phi^+, \phi^-)\|_{\mathcal{D}^s}, \frac{1}{5} \sup_M \|A_M^+\|_{Mod_M^+}, \frac{1}{5} \sup_N \|A_N^-\|_{Mod_N^-}, \frac{1}{5} \sup_{M \sim_\delta N} \|B_{M,N}\|_{C_u^s C_v^s}, \|\psi\|_{C_u^r C_v^r} \right) \leq \theta.$$

The $\frac{1}{5}$ -factors in Hypothesis 3.8 are included because of the $\theta \delta_m^a$ and $\theta \delta_n^a$ terms in (3.39) and (3.40) below. The following notation, which describes the different components in our Ansatz, is used to organize the case-analysis in the null-form estimates (Sect. 5).

Definition 3.9. (Types) Let $\zeta : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ be one of the four functions in our Ansatz. Then, we define Type $\zeta \in \{(+), (-), (+-), (S)\}$ by

$$\text{Type } \zeta := \begin{cases} (+) & \text{if } \zeta = A_M^+ \phi_M^+ \text{ for some } M, \\ (-) & \text{if } \zeta = A_N^- \phi_N^- \text{ for some } N, \\ (+-) & \text{if } \zeta = B_{M,N} \phi_M^+ \phi_N^-, \text{ for some } M \sim_\delta N, \\ (S) & \text{if } \zeta = \psi. \end{cases} \quad (3.21)$$

After proving several multi-linear estimates below, we will eventually need to sum over the dyadic frequency scales. This summation will be possible since all relevant estimates will exhibit a small gain in the frequency-scale exponents. To track this gain with minimal notational effort, we make the following definition.

Definition 3.10. (Gains) Let η be as in (2.4) and let $\zeta : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ be one of the four functions in our Ansatz. Then, we define the gain corresponding to ζ as

$$\mathcal{G}(\zeta) := \begin{cases} M^{-\eta} & \text{if } \zeta = A_M^+ \phi_M^+, \\ N^{-\eta} & \text{if } \zeta = A_N^- \phi_N^-, \\ (MN)^{-\eta} & \text{if } \zeta = B_{M,N} \phi_M^+ \phi_N^-, \\ 1 & \text{if } \zeta = \psi. \end{cases} \quad (3.22)$$

We now prove that the Ansatz (3.12) and Hypothesis 3.8 are sufficient to control the $C_t^0 C_x^s$ and $C_t^1 C_x^{s-1}$ -norms of ϕ , which will be used in the proof of Theorem 1.4.

Lemma 3.11. (Regularity in space-time) *Let Hypothesis 3.8 be satisfied. Then, it holds that*

$$\|\phi\|_{C_t^0 C_x^s(\mathbb{R} \times \mathbb{R})} + \|\phi\|_{C_t^1 C_x^{s-1}(\mathbb{R} \times \mathbb{R})} \lesssim \theta. \quad (3.23)$$

Furthermore, ϕ depends Lipschitz continuously on the linear waves ϕ^+ and ϕ^- , the modulations A^+ , A^- , and B , and the nonlinear remainder ψ with respect to the natural norm on $\mathcal{D}^s \times \text{Mod} \times C_u^r C_v^r$.

We remark that the proof of Lemma 3.11 is the only part of our argument which requires the supremum over $t \in \mathbb{R}$ in the \mathcal{D}^s -norm.

Proof. We only prove the estimate (3.23), since the Lipschitz-continuous dependence follows from a minor variant of the argument. The desired estimate for the first, second, and fourth summand in (3.12) follows directly from Corollary 2.22. As a result, it remains to prove that

$$\|B_{M,N,mn}^k(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^0 C_x^s(\mathbb{R} \times \mathbb{R})} \lesssim M^{-\eta} \theta^3, \quad (3.24)$$

$$\|B_{M,N,mn}^k(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^1 C_x^{s-1}(\mathbb{R} \times \mathbb{R})} \lesssim M^{-\eta} \theta^3. \quad (3.25)$$

In order to prove the first inequality (3.24), we estimate

$$\begin{aligned} & \|B_{M,N,mn}^k(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^0 C_x^s(\mathbb{R} \times \mathbb{R})} \\ & \lesssim \max(M, N)^s \|B_{M,N,mn}^k(u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v)\|_{L_u^\infty L_v^\infty} \\ & \lesssim \max(M, N)^s (MN)^{-s} \theta^3. \end{aligned}$$

Since $M \sim_\delta N$, this contribution is acceptable. We now turn to the second inequality (3.25) and decompose

$$\begin{aligned} & \partial_t \left(B_{M,N,mn}^k(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t) \right) \\ & = (\partial_v - \partial_u) (B_{M,N,mn}^k)(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t) \end{aligned} \quad (3.26)$$

$$+ B_{M,N,mn}^k(x-t, x+t) \partial_u \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t) \quad (3.27)$$

$$+ B_{M,N,mn}^k(x-t, x+t) \phi_M^{+,m}(x-t) \partial_v \phi_N^{-,n}(x+t). \quad (3.28)$$

For the first summand (3.26), we have that

$$\begin{aligned} & \|(\partial_v - \partial_u) (B_{M,N,mn}^k)(x-t, x+t) \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^0 C_x^{s-1}} \\ & \lesssim \|(\partial_v - \partial_u) (B_{M,N,mn}^k)(u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v)\|_{L_u^\infty L_v^\infty} \lesssim \max(M, N)^{1-s} (MN)^{-s} \theta^3. \end{aligned}$$

Since $M \sim_\delta N$ and $1 - 3s \approx -1/2$, this contribution is acceptable. For the second summand, we obtain from the definition of the \mathcal{D}^s -norm that

$$\begin{aligned} & \|B_{M,N,mn}^k(x-t, x+t) \partial_u \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^0 C_x^{s-1}} \\ & \lesssim \|B_{M,N,mn}^k(x-t, x+t)\|_{C_t^0 C_x^{1-s+\eta}} \|\partial_u \phi_M^{+,m}(x-t) \phi_N^{-,n}(x+t)\|_{C_t^0 C_x^{s-1}} \\ & \lesssim \max(M, N)^{1-2s+\eta} M^{-s} \theta^3. \end{aligned}$$

As in our estimate of (3.26), this contribution is acceptable. Since the estimate for the third summand (3.28) is similar, this completes the proof. \square

3.4. The modulation and forced wave maps equations. We now state the equations for the unknowns A^+ , A^- , B , and ψ . The equations are split into a set of modulation equations for A^+ , A^- , and B and a forced wave maps equation for ψ . Before we can define the modulation equations, we need to define the auxiliary matrices F^+ , F^- , and G , which appear as driving forces in the modulation equations.

Definition 3.12. (The forcing terms F^+ , F^- , and G) Let $1 \leq a, m, n \leq D$ and let M and N be dyadic frequency scales. Then, we define F^+ by

$$F_{M,m}^{+,a} = \sum_{N \leq M^{1-\delta}} A_{M,m}^{+,i} \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \quad (3.29)$$

$$+ \sum_{\substack{\text{Type } \zeta = \\ (+), (+-), (S)}} A_{M,m}^{+,i} \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v \zeta^j \quad (3.30)$$

$$+ \sum_{N \sim_\delta M} (\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^v \partial_v \phi_N^{-,n} \quad (3.31)$$

$$+ \sum_{N \sim_\delta M} \mathbb{I}_{ij}^{\diamond,a}(\phi) B_{M,N,mn}^j (\phi_N^{-,n} \odot_\sigma^v \partial_v \psi^j). \quad (3.32)$$

Similarly, we define F^- by

$$F_{N,n}^{-,a} = \sum_{M \leq N^{1-\delta}} A_{N,n}^{-,j} \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \quad (3.33)$$

$$+ \sum_{\substack{\text{Type } \zeta = \\ (-), (+-), (S)}} A_{N,n}^{-,j} \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u \zeta^i \quad (3.34)$$

$$+ \sum_{M \sim_\delta N} (\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^u \partial_u \phi_M^{+,m} \quad (3.35)$$

$$+ \sum_{M \sim_\delta N} \mathbb{I}_{ij}^{\diamond,a}(\phi) B_{M,N,mn}^j (\phi_M^{+,m} \odot_\sigma^u \partial_u \psi^i). \quad (3.36)$$

Finally, we define G by

$$G_{M,N,mn}^a = -P_{\leq M^{1-\sigma}}^u P_{\leq N^{1-\sigma}}^v \left(\mathbb{I}_{ij}^a(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right). \quad (3.37)$$

Instead of G , we could have denoted the right-hand side of (3.37) by $F^{+,-}$. Since the modulation equation involving G is quite different from the modulation equations involving F^+ and F^- (see Definition 3.13), we choose our notation to reflect this difference.

For notational convenience, we also define

$$G_{\chi,M,N,mn}^a = -P_{\leq M^{1-\sigma}}^u P_{\leq N^{1-\sigma}}^v \left(\chi^+ \chi^- \mathbb{I}_{ij}^a(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right). \quad (3.38)$$

While the precise definitions of F^+ , F^- , and G are quite complicated, their form will be motivated in Sect. 8. For now, they are simply convenient notation for stating the modulation equations.

Definition 3.13. (*Modulation equations*) The modulation equations for A^+ , A^- , and B are given by

$$A_{M,m}^{+,a}(u, v) = \theta \delta_m^a - P_{\leq M^{1-\sigma}}^u \chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v'), \quad (3.39)$$

$$A_{N,n}^{-,a}(u, v) = \theta \delta_n^a + P_{\leq N^{1-\sigma}}^v \chi^+ \chi^- \int_u^v du' P_{\leq N^{1-\sigma}}^v F_{N,n}^{-,a}(u', v), \quad (3.40)$$

$$B_{M,N,mn}^b(u, v) = 1\{M \sim_\delta N\} G_{\chi, M, N, mn}^b(u, v), \quad (3.41)$$

for all $1 \leq a, b, m, n \leq D$ and frequency scales M and N .

As mentioned before, we will solve the modulation equations using the para-controlled approach of [GIP15]. However, the para-controlled structure of the modulations will only be used in Sect. 8. In the PDE-analysis, which spans the majority of the paper, we will only utilize the following well-posedness result for the modulation equations.

Proposition 3.14. (Local well-posedness of the modulation equations) *Let $\phi^+, \phi^- : \mathbb{R} \rightarrow \mathbb{R}^D$ and $\psi : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ satisfy*

$$\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta \quad \text{and} \quad \|\psi\|_{C_u^r C_v^r} \leq \theta. \quad (3.42)$$

Then, there exists a solution (A^+, A^-, B) of the system of modulation equations (3.39), (3.40), and (3.41) satisfying

$$\|A_M^+ - \theta \text{Id}\|_{\text{Mod}_M^+}, \|A_N^- - \theta \text{Id}\|_{\text{Mod}_N^-}, \|B_{M,N}\|_{C_u^s C_v^s} \leq \theta \quad (3.43)$$

for all frequency-scales M and N . Furthermore, the modulations (A^+, A^-, B) depend Lipschitz continuously on $(\phi_0(0), \phi^+, \phi^-, \psi)$ in $\mathcal{M} \times \mathcal{D}^s \times C_u^r C_v^r$ (in the sense of Definition 3.3).

To emphasize the dependence on ψ , we denote the modulations from Proposition 3.14 by

$$A^+[\psi] = (A_{M,m}^{+,i}[\psi]), \quad A^-[\psi] = (A_{N,n}^{-,j}[\psi]), \quad \text{and} \quad B[\psi] = (B_{M,N,mn}^k[\psi]).$$

Similarly, we write

$$\begin{aligned} \phi^i[\psi](u, v) &= \sum_M A_{M,m}^{+,i}[\psi](u, v) \phi_M^{+,m}(u) + \sum_N A_{N,n}^{-,i}[\psi](u, v) \phi_N^{-,n}(v) \\ &\quad + \sum_{\substack{M, N: \\ M \sim_\delta N}} B_{M,N,mn}^i[\psi](u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v) + \psi^i(u, v). \end{aligned} \quad (3.44)$$

While $A^+[\psi]$, $A^-[\psi]$, $B[\psi]$, and $\phi[\psi]$ also depend on the initial position $\phi_0(0)$ and the linear waves ϕ^+ and ϕ^- , this dependence has been neglected from our notation. It now remains to state the (forced) wave maps equation for the nonlinear remainder ψ , which is determined by the wave maps equation (3.6) and our Ansatz.

Definition 3.15. (*Forced wave maps equation for ψ*) The forced wave maps equation for the nonlinear remainder ψ is given by

$$\psi^k = \theta \phi^{+,k} + \theta \phi^{-,k} - \chi^+ \chi^- \text{Duh} \left[\Pi_{ij}^{\diamond, k}(\phi[\psi]) \partial_u \phi^i[\psi] \partial_v \phi^j[\psi] \right] - (\phi^k[\psi] - \psi^k). \quad (3.45)$$

Due to the modulation equations for A^+ , A^- , and B , there are significant cancellations between the Duhamel integral and the difference $\phi^k[\psi] - \psi^k$ in (3.45). The main result on the forced wave maps equation (3.45), which utilizes these cancellations, is the following proposition.

Proposition 3.16. (Local well-posedness of the forced wave maps equation) *Let $\phi^+, \phi^- : \mathbb{R} \rightarrow \mathbb{R}^D$ satisfy $\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta$. Then, there exists a solution $\psi \in C_u^r C_v^r$ of (3.45) satisfying $\|\psi\|_{C_u^r C_v^r} \leq \theta$. Furthermore, ψ depends Lipschitz continuously on $(\phi_0(0), \phi^+, \phi^-)$ in $\mathcal{M} \times \mathcal{D}^s$ (in the sense of Definition 3.3).*

The local well-posedness of the forced wave maps equation will be proved in Sect. 9 below.

4. Prototypical Estimates

In this section, we prove several prototypical estimates. In contrast to the estimates in Sect. 2, the following prototypical estimates are specific to our Ansatz. They will then be used repeatedly in our analysis of the null-form $\partial_u \phi^i \partial_v \phi^j$ (in Sect. 5) or the full nonlinearity $\mathbb{I}_{ij}^k(\phi) \partial_u \phi^i \partial_v \phi^j$ (in Sect. 6). In Sect. 4.1, we estimate terms containing derivatives on the (low-frequency) modulations A^+ , A^- , or B . In the following, such derivatives are sometimes called “favorable”. In Sect. 4.2, we prove prototypical bilinear and trilinear estimates. In Sect. 4.3, we isolate an unfortunate resonance between the initial data ϕ^\pm and the nonlinear remainder ψ . In Sect. 4.4, we show that the types $(+)$ and $(+-)$ can sometimes be treated in a unified fashion.

4.1. Favorable derivatives. We now estimate favorable terms, i.e., terms with derivatives on the (low-frequency) modulations A^+ , A^- , and B . With the following lemma, the favorable terms can essentially be treated as smooth remainders.

Lemma 4.1. (Favorable derivatives)

$(+)$: For all $M \geq 1$, it holds that

$$\|\partial_u A_M^+ \phi_M^+\|_{C_u^{r-1} C_v^s} \lesssim M^{r-2s} \|A_M^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s}, \quad (4.1)$$

$$\|\partial_u A_M^+ \phi_M^+\|_{C_u^{r-1} C_v^r} \lesssim M^{2r-3s} \|A_M^+\|_{Mod_M^+} \|\phi^+\|_{C_u^s}. \quad (4.2)$$

$(+-)$: For all M_1, M_2 satisfying $M_1 \sim_\delta M_2$, it holds that

$$\begin{aligned} & \|\partial_u B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{C_u^{r-1} C_v^r} + \|\partial_v B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{C_u^r C_v^{r-1}} \\ & \lesssim \max(M_1, M_2)^{2r-3s} \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|\phi^-\|_{C_v^s}. \end{aligned} \quad (4.3)$$

In addition, we have the modified estimate

$$\begin{aligned} & \|\partial_u B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{C_u^{r-1} C_v^{1-r'}} + \|\partial_v B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{C_u^{1-r'} C_v^{r-1}} \\ & \lesssim \max(M_1, M_2)^{1-3s} \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|\phi^-\|_{C_v^s}. \end{aligned} \quad (4.4)$$

Proof. We treat the two cases separately.

(+)-estimate: Since A_M^+ is supported on u -frequencies $\lesssim M^{1-\sigma}$, it holds that

$$\begin{aligned} \|\partial_u A_M^+ \phi_M^+\|_{C_u^{r-1} C_v^s} &\lesssim M^{r-1} \|\partial_u A_M^+ \phi_M^+\|_{L_u^\infty C_v^s} \\ &\lesssim M^{r-1} \left(\sum_{M_1 \lesssim M^{1-\sigma}} \|\partial_u P_{M_1}^u A_M^+\|_{L_u^\infty C_v^s} \right) \|\phi_M^+\|_{L_u^\infty} \\ &\lesssim M^{r-1-s} \left(\sum_{M_1 \lesssim M^{1-\sigma}} M_1^{1-s} \right) \|A_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s} \\ &\lesssim M^{r-2s} \|A_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s}. \end{aligned}$$

This completes the proof of (4.1). The proof of (4.2) is similar and the only additional ingredient is Corollary (3.7), which yields

$$\|A_M^+\|_{C_u^s C_v^r} \lesssim M^{(1-\delta)(r-s)} \|A_M^+\|_{Mod_M^+} \lesssim M^{r-s} \|A_M^+\|_{Mod_M^+}.$$

(+-)-estimate: We only prove (4.3), since the proof of (4.4) is essentially identical. Due to the symmetry in u and v , it suffices to bound the first summand. Since B_{M_1, M_2} is supported on u -frequencies $\lesssim M_1^{1-\sigma}$ and v -frequencies $\lesssim M_2^{1-\sigma}$, we have that

$$\begin{aligned} \|\partial_u B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{C_u^{r-1} C_v^r} &\lesssim M_1^{r-1} M_2^r \|\partial_u B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-\|_{L_{u,v}^\infty} \\ &\lesssim M_1^{r-1} M_2^r \|\partial_u B_{M_1, M_2}\|_{L_{u,v}^\infty} \|\phi_{M_1}^+\|_{C^0} \|\phi_{M_2}^-\|_{C^0} \\ &\lesssim M_1^{r-1+(1-\sigma)(1-s)-s} M_2^{r-s} \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi^+\|_{C^s} \|\phi^-\|_{C^s}. \end{aligned}$$

Since $M_1 \sim_\delta M_2$ and $\sigma = 100\delta$, this is acceptable. \square

As an application of Lemma 4.1 and our commutator estimates, we obtain the following corollary.

Corollary 4.2. *For all $M \geq 1$ and $\zeta \in C_u^{1-r'} C_v^{r-1}$, it holds that*

$$\begin{aligned} &\left\| \zeta \otimes_\sigma^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) - (A_{M,m}^{+,i} \zeta) \otimes_\sigma^u \partial_u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim M^{r-2s+\sigma} \|P_{\leq M}^u \zeta\|_{C_u^s C_v^{r-1}} \|A_{M,m}^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s}. \end{aligned} \quad (4.5)$$

Proof. Since $A_{M,m}^{+,i}$ is supported on u -frequencies $\lesssim M^{1-\sigma}$, we have that

$$P_{>M}^u \zeta \otimes_\sigma^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) = (A_{M,m}^{+,i} P_{>M}^u \zeta) \otimes_\sigma^u \partial_u \phi_M^{+,m} = 0.$$

Thus, we can replace ζ by $P_{\leq M}^u \zeta$, which we now omit from our notation. Using Lemma 2.14, we obtain that

$$\begin{aligned} &\left\| \zeta \otimes_\sigma^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) - P_{\leq M^{1-\sigma}}^u \zeta \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ &+ \left\| (A_{M,m}^{+,i} \zeta) \otimes_\sigma^u \partial_u \phi_M^{+,m} - P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \zeta) \partial_u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim M^{r-2s+\sigma} \|\zeta\|_{C_u^s C_v^{r-1}} \|A_{M,m}^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s}, \end{aligned}$$

which is an acceptable contribution to (4.5). Thus, it remains to prove that

$$\begin{aligned} & \|P_{\leq M^{1-\sigma}}^u \zeta \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) - P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \zeta) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M^{r-2s+\sigma} \|\zeta\|_{C_u^s C_v^{r-1}} \|A_{M,m}^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s}. \end{aligned}$$

To this end, we decompose

$$\begin{aligned} & P_{\leq M^{1-\sigma}}^u \zeta \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) - P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \zeta) \partial_u \phi_M^{+,m} \\ & = P_{\leq M^{1-\sigma}}^u \zeta \partial_u A_{M,m}^{+,i} \phi_M^{+,m} + [A_{M,m}^{+,i}, P_{\leq M^{1-\sigma}}^u](\zeta) \partial_u \phi_M^{+,m}. \end{aligned} \quad (4.6)$$

The first summand in (4.6) can be estimated using the bilinear estimate (Proposition 2.5) and Lemma 4.1. Using frequency-support considerations and the commutator estimate (Lemma 2.10), the second summand in (4.6) can be estimated by

$$\begin{aligned} & \| [A_{M,m}^{+,i}, P_{\leq M^{1-\sigma}}^u](\zeta) \partial_u \phi_M^{+,m} \|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M^{r-1-\eta} \| [A_{M,m}^{+,i}, P_{\leq M^{1-\sigma}}^u](\zeta) \partial_u \phi_M^{+,m} \|_{C_u^\eta C_v^{r-1}} \\ & \lesssim M^{r-1-\eta} \| [A_{M,m}^{+,i}, P_{\leq M^{1-\sigma}}^u](\zeta) \|_{C_u^\eta C_v^{r-1}} \| \partial_u \phi_M^{+,m} \|_{C_u^\eta} \\ & \lesssim M^{r-1-\eta} M^{(1-\sigma)(\eta-s)} M^{\eta-s} \|\zeta\|_{C_u^s C_v^{r-1}} \|A_{M,m}^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s}. \end{aligned}$$

Since $\eta \ll \sigma$, this is an acceptable contribution to (4.5). \square

4.2. Prototypical bilinear and trilinear estimates. The next lemma essentially addresses the null form estimates in the (+)-(S) and (-)-(S) cases (see Sect. 5). However, the same estimates can also be used to treat certain error terms and therefore are placed in this central location.

Lemma 4.3. (Prototypical bilinear estimates) *Let $1 \leq i \leq D$ and let M and N be dyadic scales. Furthermore, let $\zeta \in C_b^\infty(\mathbb{R}_{u,v}^{1+1})$.*

(+)-(S): *It holds that*

$$\|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \zeta\|_{C_u^{r-1} C_v^{r-1}} \lesssim M^{r-s} \|A_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s} \|\zeta\|_{C_u^r C_v^{r-1}}. \quad (4.7)$$

By further restricting to the low \times high-interactions in the u -variable, we have the improved estimate

$$\begin{aligned} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \otimes_\sigma^u \zeta\|_{C_u^{r-1} C_v^{r-1}} & \lesssim M^{1-r-s+\eta} \|A_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s} \|\zeta\|_{C_u^r C_v^{r-1}} \\ & \lesssim M^{-\eta} \|A_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s} \|\zeta\|_{C_u^r C_v^{r-1}}. \end{aligned} \quad (4.8)$$

(-)-(S): *It holds that*

$$\|\partial_u (A_{N,n}^{-,i} \phi_N^{-,n}) \zeta\|_{C_u^{r-1} C_v^{r-1}} \lesssim N^{-\eta} \|A_N^-\|_{Mod_N^-} \|\phi_N^-\|_{C_v^s} \|\zeta\|_{C_u^{1-r'} C_v^{r-1}}. \quad (4.9)$$

Proof of Lemma 4.3: We treat the two cases separately.

(+)-(S)-estimate: Using the bilinear estimate (Proposition 2.5), we obtain that

$$\begin{aligned} \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\zeta\|_{C_u^{r-1}C_v^{r-1}} &\lesssim \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r-1}C_v^s}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ &\lesssim M^{r-s}\|A_M^+\|_{C_u^sC_v^s}\|\phi^+\|_{C_u^s}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}}. \end{aligned} \quad (4.10)$$

This proves the first estimate (4.7). Using the low \times high-improvement of the bilinear estimate, i.e., Proposition 2.5.(ii), we obtain that

$$\begin{aligned} \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\otimes_\sigma^u\zeta\|_{C_u^{r-1}C_v^{r-1}} &\lesssim \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r'}C_v^s}\|\zeta\|_{C_u^rC_v^{r-1}} \\ &\lesssim M^{1-s-r'}\|A_M^+\|_{C_u^sC_v^s}\|\phi_M^+\|_{C_u^s}\|\zeta\|_{C_u^rC_v^{r-1}}. \end{aligned}$$

(-)-(S)-estimate: We decompose

$$\partial_u(A_{N,n}^{-,i}\phi_N^{-,n})\zeta = (\partial_u A_{N,n}^{-,i}\phi_N^{-,n})\odot^v\zeta + (\partial_u A_{N,n}^{-,i}\phi_N^{-,n})\odot^v\zeta.$$

Using the non-resonant improvement in the bilinear estimate, i.e., Proposition 2.5.(iii), we have that

$$\begin{aligned} \|(\partial_u A_{N,n}^{-,i}\phi_N^{-,n})\odot^v\zeta\|_{C_u^{r-1}C_v^{r-1}} &\lesssim \|\partial_u A_{N,n}^{-,i}\phi_N^{-,n}\|_{C_u^{r-1}C_v^\eta}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ &\lesssim \|\partial_u A_{N,n}^{-,i}\|_{C_u^{r-1}C_v^s}\|\phi_N^{-,n}\|_{C_v^\eta}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ &\lesssim M^{r-2s+\eta}\|A_M^-\|_{Mod_M}\|\phi^-\|_{C_v^s}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}}. \end{aligned}$$

Since $r - 2s \approx -1/4$, this term is acceptable. Using the resonant improvement of the bilinear estimate, i.e., Proposition 2.5.(iv), and Corollary 3.7, it holds that

$$\begin{aligned} \|(\partial_u A_{N,n}^{-,i}\phi_N^{-,n})\odot^v\zeta\|_{C_u^{r-1}C_v^{r-1}} &\lesssim \|\partial_u A_{N,n}^{-,i}\phi_N^{-,n}\|_{C_u^{r-1}C_v^{r'}}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ &\lesssim N^{-s-r'}\|A_{N,n}^{-,i}\|_{C_u^rC_v^s}\|\phi_N^{-,n}\|_{C_v^s}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ &\lesssim N^{(1-\delta)(r-s)-s-r'}\|A_N^-\|_{Mod_N}\|\phi_N^-\|_{C_v^s}\|\zeta\|_{C_u^{1-r'}C_v^{r-1}}. \end{aligned}$$

Since

$$(1 - \delta)(r - s) - s - r' = 1 - 2s + \eta - \delta(r - s) \leq -\delta/10,$$

this is acceptable. \square

We now present a prototypical trilinear estimate in the initial data ϕ^\pm .

Lemma 4.4 (Prototypical trilinear estimate). *Let K , M , and N be dyadic scales and let $1 \leq k, m, n \leq D$. Then, it holds that*

$$\|\phi_K^{+,k}(u)\partial_u\phi_M^{+,m}(u)\partial_v\phi_N^{+,n}(v)\|_{C_u^{r-1}C_v^{r-1}} \lesssim K^{-s}M^{r-s}N^{r-s} \cdot \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^3. \quad (4.11)$$

In particular, the condition $K \geq \max(M, N)^{1-\sigma}$ implies the stronger estimate

$$\|\phi_K^{+,k}(u)\partial_u\phi_M^{+,m}(u)\partial_v\phi_N^{+,n}(v)\|_{C_u^{r-1}C_v^{r-1}} \lesssim (KMN)^{-\eta} \cdot \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^3. \quad (4.12)$$

A similar estimate also holds with $\phi_K^+(u)$ replaced by $\phi_K^-(v)$.

As will be clear from the proof, Lemma 4.4 easily follows from the definition of \mathcal{D}^s . Despite its simplicity, the estimate is essential for our analysis. Most importantly, it shows that our Ansatz does not need to include an explicit cubic term in ϕ^\pm .

Proof. Using Lemma 3.4, it holds that

$$\begin{aligned} \|\phi_K^{+,k} \partial_u \phi_M^{+,m} \partial_v \phi_N^{+,n}\|_{C_u^{r-1} C_v^{r-1}} &= \|\phi_K^{+,k} \partial_u \phi_M^{+,m}\|_{C_u^{r-1}} \|\partial_v \phi_N^{+,n}\|_{C_v^{r-1}} \\ &\lesssim K^{-s} M^{r-s} N^{r-s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^3. \end{aligned}$$

This implies the first estimate. If $K \geq \max(M, N)^{1-\sigma}$, the desired second estimate follows directly from our parameter conditions (2.2) and (2.4). \square

4.3. Unfortunate resonances. In this section, we describe and estimate frequency-resonances between the initial data ϕ^\pm and a smoother function ζ . These frequency-resonances between ϕ^\pm and ζ are quite different from the frequency-resonances between ϕ^\pm and itself. Using only the regularity assumption $\phi^\pm \in C_x^s$, the resonant product $\phi^{\pm,i} \odot \partial_x \phi^{\pm,j}$ is ill-defined. For our geometric random data, however, the resonant portion has been controlled using probabilistic arguments (see Sect. 2.6). In fact, Proposition 2.32 and Proposition 2.33 yield significant decay of the high \times high-product in the high frequency scale. In contrast, the resonant portion $\phi^{-,k} \odot \partial_v \zeta^n$ is well-defined using only the regularity assumptions $\phi^- \in C_v^s$ and $\zeta \in C_u^s C_v^r$. However, it exhibits worse decay in the high frequency scale than $\phi^{\pm,i} \odot \partial_x \phi^{\pm,j}$. The resulting contribution to the nonlinearity is a δ -power away from $C_u^{r-1} C_v^{r-1}$ (see Proposition 6.6) and therefore has to be absorbed in our Ansatz. While this does not create any serious (conceptual) difficulties, it feels unfortunate to barely miss $C_u^{r-1} C_v^{r-1}$, which is why we call this term the unfortunate resonance.

Lemma 4.5 (Unfortunate resonances). *Let $1 \leq k, n \leq D$, let K be a dyadic scale, and let $\zeta \in C_u^s C_v^r$. The unfortunate resonance satisfies*

$$\|\phi_K^{-,k} \odot_\sigma^v \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} \lesssim K^{1-s-r} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r}. \quad (4.13)$$

After eliminating the unfortunate resonances, the product obeys the better estimate

$$\|\phi_K^{-,k} \partial_v \zeta^n - \phi_K^{-,k} \odot_\sigma^v \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} \lesssim K^{-(r-s)-\sigma/6} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r}. \quad (4.14)$$

In total, the product satisfies the combined estimate

$$\|\phi_K^{-,k} \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} \lesssim K^{1-s-r} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r}. \quad (4.15)$$

Remark 4.6. The gain in (4.14) is a result of the v -frequency of $\phi_K^{-,k} \partial_v \zeta^n - \phi_K^{-,k} \odot_\sigma^v \partial_v \zeta^n$, which is bounded below by K^σ . After adjusting the norm on the left-hand side in (4.13) or (4.14), a similar gain is possible by inserting frequency-projections in the u -variable. For instance, it holds that

$$\begin{aligned} \|P_{\geq K^\sigma}^u (\phi_K^{-,k} \partial_v \zeta^n)\|_{C_u^{1-r'} C_v^{r-1}} &\lesssim K^{\sigma(1-r'-s)} \|\phi_K^{-,k} \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} \\ &\lesssim K^{-(r-s)-\sigma/6} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r}. \end{aligned}$$

Proof. We first prove (4.13), which is the easier estimate. It holds that

$$\begin{aligned} \|\phi_K^{-,k} \downarrow_{\sigma}^v \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} &\lesssim \|\phi_K^{-,k} \downarrow_{\sigma}^v \partial_v \tilde{P}_K^v \zeta^n\|_{C_u^s C_v^{r-1}} \\ &\lesssim \|\phi_K^{-,k}\|_{L_v^\infty} \|\partial_v \tilde{P}_K^v \zeta^n\|_{L_v^\infty C_u^s} \\ &\lesssim K^{1-s-r} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^0}. \end{aligned}$$

We now turn to the non-resonant estimate (4.14). To this end, we decompose

$$\phi_K^{-,k} \partial_v \zeta^n - \phi_K^{-,k} \downarrow_{\sigma}^v \partial_v \zeta^n = \phi_K^{-,k} \oslash^v \partial_v \zeta^n \quad (4.16)$$

$$+ \phi_K^{-,k} \ominus^v \partial_v \zeta^n - \phi_K^{-,k} \downarrow_{\sigma}^v \partial_v \zeta^n. \quad (4.17)$$

The contribution of (4.16) can be estimated using Proposition 2.5.(iii), which yields

$$\begin{aligned} \|\phi_K^{-,k} \oslash^v \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} &\lesssim \|\phi_K^{-,k}\|_{C_v^\eta} \|\partial_v \zeta\|_{C_u^s C_v^{r-1}} \\ &\lesssim K^{-(s-n)} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^0}. \end{aligned}$$

Since $-s \approx -1/2$ and $-(r-s) \approx -1/4$, this term is acceptable. The contribution of (4.17) is estimated by

$$\begin{aligned} \|\phi_K^{-,k} \ominus^v \partial_v \zeta^n - \phi_K^{-,k} \downarrow_{\sigma}^v \partial_v \zeta^n\|_{C_u^s C_v^{r-1}} &\lesssim K^{\sigma(r-1)} \|\phi_K^{-,k} \ominus^v \partial_v \zeta^n\|_{C_u^s C_v^0} \\ &\lesssim K^{\sigma(r-1)} \|\phi_K^{-,k}\|_{L_v^\infty} \|\tilde{P}_K^v \partial_v \zeta^n\|_{L_v^\infty C_u^s} \\ &\lesssim K^{\sigma(r-1)-s-(r-1)} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^{r-1}}. \end{aligned}$$

Since

$$\sigma(r-1) - s - (r-1) = -(r-s) + 1 - 2s - \sigma(1-r) \leq -(r-s) - \sigma/6,$$

this term is acceptable. The combined estimate (4.15) follows directly from (4.13), (4.14), and the triangle inequality. \square

4.4. Generalizations of (+) and (+-). In several (but not all) estimates below, it is possible to treat the (+) and (+-) terms simultaneously, which makes the case analysis more efficient. This essentially means that only the right-moving component of the wave is relevant. The imaginative reader may think of the next lemma as a violation of common rules for pedestrians, since it allows us to look in only one direction before safely crossing a street.

Lemma 4.7 (“Looking in only one direction”). *Let $1 \leq i \leq D$, M, M_1, M_2 be dyadic scales satisfying $M = M_1 \sim_\delta M_2$, and let*

$$\begin{aligned} \zeta^{+,i}(u, v) &= A_{M,m}^{+,i}(u, v) \phi_M^{+,m}(u) \quad \text{or} \\ \zeta^{+,i}(u, v) &= B_{M_1, M_2, m_1 m_2}^i(u, v) \phi_{M_1}^{+,m_1}(u) \phi_{M_2}^{-,m_2}(v). \end{aligned}$$

Then, it holds that

$$\zeta^{+,i} = E_{M,m}^{+,i}(u, v) \phi_M^{+,m}(u),$$

where $E_M^{+,i}$ satisfies either the estimates

$$\begin{aligned} \|E_M^{+,i}\|_{C_u^s C_v^s} &\lesssim \|A_M^+\|_{C_u^s C_v^s}, \\ \|E_M^{+,i}\|_{C_u^s C_v^r} &\lesssim M^{(1-\delta)(r-s)} \|A_M^+\|_{Mod_N^+}, \end{aligned} \quad (4.18)$$

or the estimates

$$\begin{aligned} \|E_M^{+,i}\|_{C_u^s C_v^s} &\lesssim \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi_{M_1}^-\|_{C_v^s}, \\ \|E_M^{+,i}\|_{C_u^s C_v^r} &\lesssim M^{(1+2\delta)(r-s)} \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi_{M_1}^-\|_{C_v^s}. \end{aligned} \quad (4.19)$$

Proof. We treat the (+) and (+−) cases separately. In the (+) case, we set $E_{M,m}^{+,i} := A_{M,m}^{+,i}$. Then, the first line in (4.18) is trivial and the second line in (4.18) follows from Corollary 3.7.

We now turn to the (+−)-case. Here, we define $E_{M,m}^{+,i} = B_{M_1, M_2, m_1, m_2}^i \phi_{M_2}^{-, m_2}$, where $M_1 = M$ and $m_1 = m$. The first line in (4.19) then follows directly from the bilinear estimate (Proposition 2.5). Due to the frequency-support conditions on B , we have that

$$\|E_{M,m}^{+,i}\|_{C_u^s C_v^r} \lesssim M_2^{r-s} \|E_{M,m}^{+,i}\|_{C_u^s C_v^s}.$$

Since $M_2 \leq M_1^{1/(1-\delta)} = M^{1/(1-\delta)}$, this yields the second line in (4.19). \square

5. Bilinear Analysis

In this section, we analyze two different bilinear expressions. In Sect. 5.1, we analyze the null-form $\partial_u \phi^i \partial_v \phi^j$, which is the most important component of the nonlinearity. In Sect. 5.2, we analyze the product $\phi^k \partial_u \phi^i$. Together with Bony's para-linearization (Lemma 2.9), this will yield sufficient information on $\Pi_{ij}^{\diamond, k}(\phi) \partial_u \phi^i$.

5.1. The null-form $\partial_u \phi^i \partial_v \phi^j$. In this section, we analyze the null-form $\partial_u \phi^i \partial_v \phi^j$. The main goal is to identify the terms in $\partial_u \phi^i \partial_v \phi^j$ which belong to $C_u^{r-1} C_v^{r-1}$, since they can be treated as smooth remainders. The rough terms which cannot be placed in $C_u^{r-1} C_v^{r-1}$ will later be absorbed in our Ansatz, i.e, the terms $A_M^+ \phi_M^+$, $A_M^- \phi_M^-$, and $B_{M_1, M_2} \phi_{M_1}^+ \phi_{M_2}^-$. Unfortunately, the argument requires a tedious case-by-case analysis, which is displayed in Fig. 2. The available estimates are illustrated through the graphical symbols \times , \otimes_σ^u , $\rightarrow \downarrow_\sigma^v$, and \checkmark , which will now be explained.

The main term (+)−(−) cannot be controlled in $C_u^{r-1} C_v^{r-1}$. While certain favorable sub-terms can be removed, it is more efficient to postpone a detailed analysis until the pre-factor $\Pi_{ij}^{\diamond, k}(\phi)$ is included. The graphical symbol \times illustrates that this term is not controlled here. The cases (+)−(+), (+)−(+−), and (+)−(S) can be controlled as long as the second argument is restricted to high u -frequencies. Thus, we can treat the \otimes_σ^u -portion of this term, which is also used as a graphical symbol. In the (+−)−(S)-term, the only problem lies in the unfortunate resonance between ϕ^- and $\partial_v \psi$. Since the unfortunate resonance is given by \downarrow_σ^v , we use $\rightarrow \downarrow_\sigma^v$ to illustrate this case. The remaining cases (on

$\partial_u \partial_v$	$(-)$	$(+)$	$(+-)$	(S)
$(+)$	\times	\lesssim_σ^u	\lesssim_σ^u	\lesssim_σ^u
$(-)$		\checkmark	\checkmark	\checkmark
$(+-)$			\checkmark	$\neg \downarrow_\sigma^v$
(S)				\checkmark

Fig. 2. Overview of cases in the null-form estimate

or above the diagonal in Fig. 2) can be controlled in $C_u^{r-1}C_v^{r-1}$. These terms will not need to be considered further and therefore deserve the symbol \checkmark .

Finally, we note that all cases below the diagonal in Fig. 2 have been left empty. The results in these cases can be obtained from the symmetry in the u and v -variables and do not need to be considered separately.

Proposition 5.1 (\times -estimate). *Let Hypothesis 3.8 be satisfied, let $1 \leq i, j \leq D$, and let M, N be frequency scales. Then, it holds that*

$$\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim M^{r-s} N^{r-s} \theta^4. \quad (5.1)$$

If $M \sim_\delta N$, we also have that

$$\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) - A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim (MN)^{-\eta} \theta^4. \quad (5.2)$$

Remark 5.2. Since r is bigger than s , the first estimate (5.1) does not provide a uniform upper bound. Nevertheless, it illustrates the gap to $C_u^{r-1}C_v^{r-1}$. The second estimate (5.2) is the first step in our analysis of the main term. A more detailed analysis, however, is postponed until Sect. 6.1, since it is more efficient to directly treat

$$\mathbb{I}_{ij}^{\diamond, k}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}.$$

Proof. We first prove (5.1). Using the bilinear estimate (Proposition 2.5) and frequency-support considerations, we have that

$$\begin{aligned} \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} &\lesssim \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \left\| \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^s C_v^{r-1}} \\ &\lesssim M^{r-s} N^{r-s} \left\| A_{M,m}^{+,i} \phi_M^{+,m} \right\|_{C_u^s C_v^s} \left\| A_{N,n}^{-,j} \phi_N^{-,n} \right\|_{C_u^s C_v^s} \\ &\lesssim M^{r-s} N^{r-s} \|A_M^+\|_{C_u^s C_v^s} \|\phi^+\|_{C^s} \|A_N^-\|_{C_u^s C_v^s} \|\phi^-\|_{C^s}. \end{aligned}$$

We now turn to (5.2). To this end, we decompose

$$\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) - A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}$$

$$= \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v A_{N,n}^{-,j} \phi_N^{-,n} \quad (5.3)$$

$$+ \partial_u A_{M,m}^{+,i} \phi_M^{+,m} \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \quad (5.4)$$

$$- \partial_u A_{M,m}^{+,i} \phi_M^{+,m} \partial_v A_{N,n}^{-,j} \phi_N^{-,n}. \quad (5.5)$$

Using the prototypical bilinear estimate (Lemma 4.3) and Lemma 4.1, the first summand (5.3) is estimated by

$$\begin{aligned} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v A_{N,n}^{-,j} \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} &\lesssim M^{r-s} \|A_M^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|\partial_v A_{N,n}^{-,j} \phi_N^{-,n}\|_{C_u^{1-r'} C_v^{r-1}} \\ &\lesssim M^{r-s} N^{r-2s} \theta^4. \end{aligned}$$

Since $M \sim_\delta N$ and $3s - 2r \gg \delta$, this term is acceptable. By symmetry, the second summand (5.4) obeys the same estimate. Finally, the third summand (5.5) can be estimated easily using the bilinear estimate (Proposition 2.5) and Lemma 4.1. \square

We now turn to the \mathbb{S}_σ^μ -cases in Fig. 2.

Proposition 5.3. (\mathbb{S}_σ^μ -estimates) *Let Hypothesis 3.8 be satisfied, let $1 \leq i, j \leq D$, and let M be a frequency scale. Furthermore, let $\zeta: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ be of type (+), (+-), or (S). Then, with $\mathcal{G}(\zeta)$ as in (3.22), it holds that*

$$\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \mathbb{S}_\sigma^\mu \partial_v \zeta^j \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim M^{-\eta} \mathcal{G}(\zeta) \theta^3. \quad (5.6)$$

Without the low \times high-para-product, we only have the weaker estimate

$$\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v \zeta^j \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim M^{r-s} \mathcal{G}(\zeta) \theta^3. \quad (5.7)$$

Proof. The case (+)-(S) has already been treated in Lemma 4.3. Thus, it remains to treat the cases (+)-(+) and (+)-(+), which can be done simultaneously. Using Lemma 4.7, we can write

$$\zeta^j(u, v) = E_{N,n}^{+,j}(u, v) \phi_N^{+,n}(u), \quad (5.8)$$

where E_N^+ satisfies either (4.18) or (4.19). In both cases, we have that

$$\|E_N^+\|_{C_u^s C_v^r} \lesssim N^{(1+2\delta)(r-s)\theta}. \quad (5.9)$$

The main step in this argument is to prove

$$\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (E_{N,n}^{+,j} \phi_N^{+,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim M^{r-s} N^{r-2s+4\delta} \mathcal{G}(\zeta) \theta^4. \quad (5.10)$$

Once (5.10) has been proven, the first estimate (5.6) follows by considering $N \gtrsim M^{1-\sigma}$ and the second estimate (5.7) follows by considering all $N \geq 1$. The additional Littlewood-Paley operators, which enter through the para-product \mathbb{S}_σ^μ , can be handled using Lemma 2.10 and Lemma 2.14.

Using the product rule, we decompose

$$\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (E_{N,n}^{+,j} \phi_N^{+,n}) = A_{M,m}^{+,i} \partial_u \phi_M^{+,m} \partial_v E_{N,n}^{+,j} \phi_N^{+,n} \quad (5.11)$$

$$+ \partial_u A_{M,m}^{+,i} \phi_M^{+,m} \partial_v E_{N,n}^{+,j} \phi_N^{+,n}. \quad (5.12)$$

We start by estimating the first summand (5.11), which is the main term. Using the multiplication estimate (Corollary 2.8), the bilinear estimate (2.5), and Hypothesis 3.8, we have that

$$\begin{aligned}
 \|A_{M,m}^{+,i} \partial_u \phi_M^{+,m} \partial_v E_{N,n}^{+,j} \phi_N^{+,n}\|_{C_u^{r-1} C_v^{r-1}} &\lesssim \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\partial_u \phi_M^{+,m} \partial_v E_{N,n}^{+,j} \phi_N^{+,n}\|_{C_u^{r-1} C_v^{r-1}} \\
 &\lesssim \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\partial_v E_{N,n}^{+,j}\|_{C_u^s C_v^{r-1}} \|\partial_u \phi_M^{+,m} \phi_N^{+,n}\|_{C_u^{r-1}} \\
 &\lesssim M^{r-s} N^{-s} \|A_M^+\|_{C_u^s C_v^s} \|E_N^+\|_{C_u^s C_v^r} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2 \\
 &\lesssim M^{r-s} N^{(1+2\delta)(r-s)-s} \theta^4.
 \end{aligned}$$

This is an acceptable contribution to (5.10). For the second summand (5.12), we have that

$$\begin{aligned}
 &\|\partial_u A_{M,m}^{+,i} \phi_M^{+,m} \partial_v E_{N,n}^{+,j} \phi_N^{+,n}\|_{C_u^{r-1} C_v^{r-1}} \\
 &\lesssim \|\partial_u A_{M,m}^{+,i}\|_{C_u^\eta C_v^\eta} \|\phi_M^{+,m}\|_{C_u^\eta} \|\partial_v E_{N,n}^{+,j}\|_{C_u^\eta C_v^{r-1}} \|\phi_N^{+,n}\|_{C_u^\eta} \\
 &\lesssim M^{1-s+\eta} M^{-s+\eta} N^{(1+2\delta)(r-s)} N^{-s+\eta} \|A_M^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|E_{N,n}^{+,j}\|_{C_u^s C_v^r} \|\phi^+\|_{C_u^s} \\
 &\lesssim M^{1-2s+2\eta} N^{r-2s+4\delta} \theta^4.
 \end{aligned}$$

This is an acceptable contribution to (5.10). \square

We now turn to the $\neg \downarrow_\sigma^v$ -case in Fig. 2.

Proposition 5.4. ($\neg \downarrow_\sigma^v$ -estimate) *Let Hypothesis 3.8 be satisfied, let $1 \leq i, j \leq D$, and let M_1, M_2 be frequency scales satisfying $M_1 \sim_\delta M_2$. Furthermore, let $\zeta \in C_u^s C_v^r$. Then, it holds that*

$$\begin{aligned}
 &\|\partial_u (B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2}) \partial_v \zeta^j - (B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j)) \ll_\sigma^u \partial_u \phi_{M_2}^{+,m_2}\|_{C_u^{r-1} C_v^{r-1}} \\
 &\lesssim (M_1 M_2)^{-\eta} \theta^3 \|\zeta\|_{C_u^s C_v^r}.
 \end{aligned} \tag{5.13}$$

Furthermore, the unfortunate resonance satisfies

$$\|B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j)\|_{C_u^s C_v^{r-1}} \lesssim M_2^{1-s-r} \theta^2 \|\zeta\|_{C_u^s C_v^r}. \tag{5.14}$$

Proof. The estimate of the unfortunate resonance (5.14) follows directly from the multiplication estimate (Corollary 2.8) and Lemma 4.5. Thus, it remains to prove (5.13). To this end, we decompose

$$\begin{aligned}
 &\partial_u (B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2}) \partial_v \zeta^j - (B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j)) \ll_\sigma^u \partial_u \phi_{M_2}^{+,m_2} \\
 &= B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \partial_v \zeta^j - \phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j) \partial_u \phi_{M_1}^{+,m_1}
 \end{aligned} \tag{5.15}$$

$$+ (B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j)) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \tag{5.16}$$

$$+ \partial_u B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2} \partial_v \zeta^j. \tag{5.17}$$

We estimate the three terms (5.15), (5.16), and (5.17) separately. Using Lemma 4.5, the first summand (5.15) is estimated by

$$\|B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \partial_v \zeta^j - \phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \zeta^j) \partial_u \phi_{M_1}^{+,m_1}\|_{C_u^{r-1} C_v^{r-1}}$$

$$\begin{aligned}
&\lesssim \|B_{M_1, M_2, m_1 m_2}^i\|_{C_u^s C_v^s} \|\phi_{M_2}^{-, m_2} \partial_v \zeta^j - \phi_{M_2}^{-, m_2} \odot_\sigma^v \partial_v \zeta^j\|_{C_u^s C_v^{r-1}} \|\partial_u \phi_{M_1}^{+, m_1}\|_{C_u^{r-1} C_v^s} \\
&\lesssim M_1^{r-s} M_2^{-(r-s)-\sigma/6} \theta^3 \|\zeta\|_{C_u^s C_v^r}.
\end{aligned}$$

Since $M_1 \sim_\delta M_2$ and $\sigma = 100\delta$, this term is acceptable. We now turn to the second summand (5.16). To this end, we first note that the tuple $(\alpha_1, \beta_1) = (s, -s + \eta)$ satisfies

$$r - 1 \leq \alpha_1 + (1 - \sigma)\beta_1 \quad \text{and} \quad \alpha_1 + \beta_1 > 0.$$

Using Proposition 2.5(ii) and Lemma 4.5, it follows that

$$\begin{aligned}
&\left\| \left(B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-, m_2} \odot_\sigma^v \partial_v \zeta^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+, m_1} \right\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \|B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-, m_2} \odot_\sigma^v \partial_v \zeta^j)\|_{C_u^s C_v^{r-1}} \|\partial_u \phi_{M_1}^{+, m_1}\|_{C_u^{-s+\eta} C_v^s} \\
&\lesssim M_1^{1-2s+\eta} M_2^{1-s-r} \|B\|_{C_u^s C_v^s} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r} \|\phi^+\|_{C_u^s} \\
&\lesssim M_1^{1-2s+\eta} M_2^{1-s-r} \theta^3 \|\zeta\|_{C_u^s C_v^r}.
\end{aligned}$$

Since $M_1 \sim_\delta M_2$ and $1 - s - r \approx -1/4$, this is acceptable. It remains to estimate the third summand (5.17). Using the bilinear estimate (Proposition 2.5) and Lemma 4.1, we have that

$$\begin{aligned}
&\|\partial_u B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \partial_v \zeta^j\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \|\partial_u B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2}\|_{C_u^{r-1} C_v^{1-r'}} \|\partial_v \zeta^j\|_{C_u^{r-1} C_v^s} \\
&\lesssim M_{\max}^{1-3s} \|B_{M_1, M_2}\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|\phi^-\|_{C_v^s} \|\zeta\|_{C_u^s C_v^r} \\
&\lesssim M_{\max}^{1-3s} \theta^3 \|\zeta\|_{C_u^s C_v^r}.
\end{aligned}$$

Since $1 - 3s \approx -1/2$, this term is acceptable. \square

We now turn to the final proposition in our case analysis, which addresses the \checkmark -terms.

Proposition 5.5. (\checkmark -estimates) *Let Hypothesis 3.8 be satisfied and let $1 \leq i, j \leq D$. Furthermore, let $\zeta_1, \zeta_2: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ satisfy*

$$\text{Type}(\zeta_1, \zeta_2) \in \{(-)-(+), (-)-(+ -), (-)-(S), (+-)-(+ -), (S)-(S)\}.$$

Then, it holds that

$$\|\partial_u \zeta_1^i \partial_v \zeta_2^j\|_{C_u^{r-1} C_v^{r-1}} \lesssim \mathcal{I}(\zeta_1) \mathcal{I}(\zeta_2) \theta^2. \quad (5.18)$$

Proof. We treat the five different cases in this proposition separately.

(-)-(+) estimate: In this case, we have that

$$\begin{aligned}
\zeta_1^i(u, v) &= A_{M,m}^{-,i}(u, v) \phi_M^{-,m}(v) \quad \text{and} \\
\zeta_2^j(u, v) &= A_{N,n}^{+,j}(u, v) \phi_N^{+,n}(u).
\end{aligned}$$

for some $M, N \geq 1$. Using the bilinear estimate (Proposition 2.5) twice and Corollary 3.7, we obtain that

$$\|\partial_u A_{M,m}^{-,i} \phi_M^{-,m} \partial_v A_{N,n}^{+,j} \phi_N^{+,n}\|_{C_u^{r-1} C_v^{r-1}}$$

$$\begin{aligned}
&\lesssim \|\phi_M^{-,m} \phi_N^{+,n}\|_{C_u^{1-r'} C_v^{1-r'}} \|\partial_u A_{M,m}^{-,i} \partial_v A_{N,n}^{+,j}\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \|\phi_M^{-,m} \phi_N^{+,n}\|_{C_u^{1-r'} C_v^{1-r'}} \|\partial_u A_{M,m}^{-,i}\|_{C_u^{r-1} C_v^s} \|\partial_v A_{N,n}^{+,j}\|_{C_u^s C_v^{r-1}} \\
&\lesssim (MN)^{1-r'-s+(1-\delta)(r-s)} \|\phi^+\|_{C^s} \|\phi^-\|_{C^s} \|A_{M,m}^-\|_{Mod_M} \|A_N^+\|_{Mod_N} \\
&\lesssim (MN)^{1-2s+\eta-\delta(r-s)} \theta^4.
\end{aligned}$$

Since

$$1 - 2s + \eta - \delta(r - s) \leq 1 - 2s + \eta - \delta/8 \leq -\delta/10,$$

this contribution is acceptable.¹¹

(-)-(+-) estimate: In this case, we have that

$$\begin{aligned}
\zeta_1^i(u, v) &= A_{M,m}^{-,i}(u, v) \phi_M^{-,m}(v) \quad \text{and} \\
\zeta_2^j(u, v) &= B_{N_1, N_2, n_1 n_2}^j(u, v) \phi_{N_1}^{+,n_1}(u) \phi_{N_2}^{-,n_2}(v)
\end{aligned}$$

for some $M, N_1, N_2 \geq 1$ satisfying $N_1 \sim_\delta N_2$. Using the product rule, we obtain

$$\partial_u \left(A_{M,m}^{-,i} \phi_M^{-,m} \right) \partial_v \left(B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+,n_1} \phi_{N_2}^{-,n_2} \right) = \partial_u A_{M,m}^{-,i} \phi_M^{-,m} B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+,n_1} \partial_v \phi_{N_2}^{-,n_2} \quad (5.19)$$

$$+ \partial_u A_{M,m}^{-,i} \phi_M^{-,m} \partial_v B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+,n_1} \phi_{N_2}^{-,n_2}. \quad (5.20)$$

We start with the first summand (5.19), which is the main term. Using the multiplication estimate (Corollary 2.8), the bilinear estimate (Proposition 2.5), and the properties of the linear waves (Corollary 3.4), we obtain that

$$\begin{aligned}
&\|\partial_u A_{M,m}^{-,i} \phi_M^{-,m} B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+,n_1} \partial_v \phi_{N_2}^{-,n_2}\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \|B_{N_1, N_2, n_1 n_2}^j\|_{C_u^s C_v^s} \|\partial_u A_{M,m}^{-,i} \phi_M^{-,m} \phi_{N_1}^{+,n_1} \partial_v \phi_{N_2}^{-,n_2}\|_{C_u^{r-1} C_v^{r-1}} \\
&\lesssim \|B_{N_1, N_2, n_1 n_2}^j\|_{C_u^s C_v^s} \|\partial_u A_{M,m}^{-,i} \phi_{N_1}^{+,n_1}\|_{C_u^\eta C_v^s} \|\phi_M^{-,m} \partial_v \phi_{N_2}^{-,n_2}\|_{C_u^\eta C_v^{r-1}} \\
&\lesssim M^{1-s+\eta} N_1^{-s+\eta} (1\{M \lesssim N_2\} N_2^{r-s} M^{-s} + 1\{M \gg N_2\} M^{r-1-s} N_2^{1-s}) \\
&\times \|A_M^+\|_{C_u^s C_v^s} \|B_{N_1, N_2}\|_{C_u^s C_v^s} \|\phi^+\|_{C^s} \|\phi^-\|_{\mathcal{D}^s}^2 \\
&\lesssim M^{1-s+\eta} N_1^{-s+\eta} (1\{M \lesssim N_2\} N_2^{r-s} M^{-s} + 1\{M \gg N_2\} M^{r-1-s} N_2^{1-s}) \theta^5.
\end{aligned}$$

Using $N_1 \sim_\delta N_2$, the pre-factor can be estimated by

$$\begin{aligned}
&M^{1-s+\eta} N_1^{-s+\eta} (1\{M \lesssim N_2\} N_2^{r-s} M^{-s} + 1\{M \gg N_2\} M^{r-1-s} N_2^{1-s}) \\
&\lesssim 1\{M \lesssim N_2\} M^{1-2s+\eta} N_2^{r-2s+\eta+\delta} + 1\{M \gg N_2\} M^{r-2s+\eta} N_2^{1-2s+\eta+\delta} \\
&\lesssim \max(M, N_1, N_2)^{r+1-4s+4\delta}.
\end{aligned}$$

¹¹ This is the shortest argument which yields an acceptable contribution. By using frequency-support considerations for A_M^- and A_N^+ , we can obtain better decay in M and N , and thus this term is less serious than our argument suggests.

Since $r + 1 - 4s \approx -1/4$, this completes our estimate of (5.19). The estimate for the minor term (5.20) follows directly from the $(-)$ -case in Lemma 4.3 and the $(+-)$ -case in Lemma 4.1.

$(-)-(S)$ estimate: This estimate directly follows from the $(-)-(S)$ -case in Lemma 4.3.

$(+-)-(+-)$ estimate: We remark that this term obeys better estimates than the $(+-)-(S)$ -term (see Proposition 5.4). The reason is that the structure of $(+-)$ allows for better bounds on certain high \times high \rightarrow low-interactions.

In this case, we have that

$$\zeta_1^i(u, v) = B_{M_1, M_2, m_1 m_2}^i(u, v) \phi_{M_1}^{+, m_1}(u) \phi_{M_2}^{-, m_2}(v) \quad \text{and}$$

$$\zeta_2^j(u, v) = B_{N_1, N_2, n_1 n_2}^j(u, v) \phi_{N_1}^{+, n_1}(u) \phi_{N_2}^{-, n_2}(v)$$

for some $M_1, M_2, N_1, N_2 \geq 1$ satisfying $M_1 \sim_\delta M_2$ and $N_1 \sim_\delta N_2$. To simplify the notation, we set $M_{\max} := \max(M_1, M_2)$ and $N_{\max} := \max(N_1, N_2)$. By symmetry, we can assume that $M_{\max} \leq N_{\max}$. Using the product rule, we have that

$$\begin{aligned} & \partial_u \left(B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \right) \partial_v \left(B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \phi_{N_2}^{-, n_2} \right) \\ &= B_{M_1, M_2, m_1 m_2}^i \partial_u \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \partial_v \phi_{N_2}^{-, n_2} \end{aligned} \quad (5.21)$$

$$+ B_{M_1, M_2, m_1 m_2}^i \partial_u \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \partial_v B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \phi_{N_2}^{-, n_2} \quad (5.22)$$

$$+ \partial_u B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \partial_v \phi_{N_2}^{-, n_2} \quad (5.23)$$

$$+ \partial_u B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \partial_v B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \phi_{N_2}^{-, n_2}. \quad (5.24)$$

We now estimate the four terms (5.21)–(5.24) separately. We begin with the main term (5.21). Using the multiplication estimate (Corollary 2.8) and Lemma 3.4, we obtain that

$$\begin{aligned} & \| B_{M_1, M_2, m_1 m_2}^i \partial_u \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \partial_v \phi_{N_2}^{-, n_2} \|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \| B_{M_1, M_2, m_1 m_2}^i \|_{C_u^s C_v^s} \| B_{N_1, N_2, n_1 n_2}^j \|_{C_u^s C_v^s} \| \partial_u \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \phi_{N_1}^{+, n_1} \partial_v \phi_{N_2}^{-, n_2} \|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M_1^{r-s} N_1^{-s} M_2^{-s} N_2^{-s} \| B_{M_1, M_2} \|_{C_u^s C_v^s} \| B_{N_1, N_2} \|_{C_u^s C_v^s} \| (\phi^+, \phi^-) \|_{\mathcal{D}^s}^4 \\ & \lesssim M_{\max}^{r-2s+\delta} N_{\max}^{r-2s+\delta} \| B_{M_1, M_2} \|_{C_u^s C_v^s} \| B_{N_1, N_2} \|_{C_u^s C_v^s} \| (\phi^+, \phi^-) \|_{\mathcal{D}^s}^4 \\ & \lesssim M_{\max}^{r-2s+\delta} N_{\max}^{r-2s+\delta} \theta^6. \end{aligned}$$

Since $r - 2s \approx -1/4$, this contribution is acceptable. We now turn to the second summand (5.22), which is relatively easy. Using the bilinear estimate (Proposition 2.5) and Lemma 3.4, we have that

$$\begin{aligned} & \| B_{M_1, M_2, m_1 m_2}^i \partial_u \phi_{M_1}^{+, m_1} \phi_{M_2}^{-, m_2} \partial_v B_{N_1, N_2, n_1 n_2}^j \phi_{N_1}^{+, n_1} \phi_{N_2}^{-, n_2} \|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \| B_{M_1, M_2, m_1 m_2}^i \|_{C_u^s C_v^s} \| \partial_u \phi_{M_1}^{+, m_1} \phi_{N_1}^{+, n_1} \|_{C_u^{r-1}} \| \partial_v B_{N_1, N_2, n_1 n_2}^j \phi_{M_2}^{-, m_2} \phi_{N_2}^{-, n_2} \|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M_1^{r-s} N_1^{-s} N_2^{1-s} M_2^{-s} N_2^{-s} \theta^6 \\ & = M_1^{r-s} M_2^{-s} N_1^{-s} N_2^{1-2s} \theta^6. \end{aligned}$$

Since $r - 2s \approx -1/4$ and $1 - 3s \approx -1/2$, this contribution is acceptable. The third summand (5.23) can be treated similarly. The fourth summand (5.24) can easily be controlled through the bilinear estimate (Proposition 2.5) and Lemma 4.1.

$(S)-(S)$ estimate: This estimate directly follows from the bilinear estimate (Proposition 2.5). \square

5.2. *The product $\phi^k \partial_u \phi^i$.* The following lemma controls $\phi^k \otimes_\sigma^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m})$ in $C_u^{r-1} C_v^{1-r'}$. Together with Bony's parilinearization (Lemma 2.9), it will be used in Sect. 6 to control $\Pi_{ij}^{\circ,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m})$.

Lemma 5.6. *Let Hypothesis 3.8 be satisfied, let $1 \leq i, k \leq D$, and let K, M be frequency scales satisfying $K \gtrsim M^{1-\sigma}$. Let $H_M^+ : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^{D \times D}$ satisfy $P_{\gg M^{1-\sigma}}^u H_M^+ = 0$. Furthermore, let $\zeta : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ be of one of the four types $(-)$, $(+)$, $(+-)$, or (S) . Then, it holds that*

$$\begin{aligned} & \left\| P_K^u (\zeta^k) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} \\ & \lesssim \left(1\{K \lesssim M\} M^{1-s-r} + 1\{K \gg M\} K^{2(r-1+\eta)} \right) \mathcal{G}(\zeta) \theta^2 \|H_M^+\|_{C_u^s C_v^s} \quad (5.25) \\ & \lesssim K^{-\eta} M^{1-s-r+\eta} \mathcal{G}(\zeta) \theta^2 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

In the individual cases, we have the following more detailed estimates, which are stated only for the (difficult) regime $M^{1-\sigma} \leq K \lesssim M$.

$(-)$: If $\zeta = A_{L,\ell}^- \phi_L^{-,\ell}$ for some $L \geq 1$, we have that

$$\begin{aligned} & \left\| P_K^u (A_{L,\ell}^{-,k} \phi_L^{-,\ell}) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} \\ & \lesssim (L^{-\eta} M^{-(r-s)-\delta/8} + L^{1-r'-s} M^{1-s-r}) \theta^3 \|H_M^+\|_{C_u^s C_v^s} \\ & \lesssim L^{-\eta} M^{1-s-r} \theta^3 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

$(+)$: If $\zeta = A_{L,\ell}^{+,k} \phi_L^{+,\ell}$ for some $L \geq 1$, then

$$\left\| P_K^u (A_{L,\ell}^{+,k} \phi_L^{+,\ell}) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \lesssim L^{-\eta} M^{r-2s+\sigma} \theta^3 \|H_M^+\|_{C_u^s C_v^s}.$$

$(+-)$: If $\zeta = B_{L_1, L_2, \ell_1 \ell_2} \phi_{L_1}^{+, \ell_1} \phi_{L_2}^{-, \ell_2}$ with $L_1 \sim_\delta L_2$, then

$$\left\| P_K^u (B_{L_1, L_2, \ell_1 \ell_2}^{k, \ell_1 \ell_2} \phi_{L_1}^{+, \ell_1} \phi_{L_2}^{-, \ell_2}) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \lesssim (L_1 L_2)^{-\eta} M^{r-2s+\sigma} \theta^4 \|H_M^+\|_{C_u^s C_v^s}.$$

(S) : If $\zeta = \psi$, then

$$\left\| P_K^u \psi^k \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) - H_{M,m}^{+,i} (P_K^u \psi^k \odot_\sigma^u \partial_u \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \lesssim M^{-(r-s)-\sigma/6} \theta^2 \|H_M^+\|_{C_u^s C_v^s}.$$

Furthermore, the resonant part satisfies

$$\left\| H_{M,m}^{+,i} (P_K^u \psi^k \odot_\sigma^u \partial_u \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \lesssim M^{1-s-r} \theta^2 \|H_M^+\|_{C_u^s C_v^s}.$$

Remark 5.7. The matrix H_M^+ will be chosen either as the modulation A_M^+ or the identity Id. By introducing the auxiliary variable H_M^+ , we are able to treat both cases simultaneously.

Proof. We split the argument into several steps. In the first three steps, we treat the case $(-)$, the two cases $(+)$ and $(+-)$ simultaneously, and the case (S) . In the final step, we combine the estimates to prove (5.25).

Step 1: The case $(-)$. We first note that

$$P_K^u(A_{L,\ell}^{-,k}\phi_L^{-,\ell}) = (P_K^u A_{L,\ell}^{-,k})\phi_L^{-,\ell}$$

We further distinguish the cases $M^{1-\sigma} \leq K \ll M$ and $K \sim M$.

Step 1.(a): $M^{1-\sigma} \leq K \ll M$. We use Proposition 2.5.(iii), which yields

$$\begin{aligned} & \left\| P_K^u A_{L,\ell}^{-,k} \phi_L^{-,\ell} \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} \\ & \lesssim \left\| P_K^u A_{L,\ell}^{-,k} \phi_L^{-,\ell} \right\|_{C_u^\eta C_v^{1-r'}} \left\| \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \\ & \lesssim K^{-s+\eta} L^{1-r'-s} M^{r-s} \|A_L^-\|_{C_u^s C_v^s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \|H_M^+\|_{C_u^s C_v^s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \\ & \lesssim K^{-s+\eta} L^{1-r'-s} M^{r-s} \theta^3 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

Since $K \gtrsim M^{1-\sigma}$, we have that

$$K^{-s+\eta} L^{1-r'-s} M^{r-s} \lesssim L^{-\eta} M^{r-2s+\sigma} \lesssim L^{-\eta} M^{-(r-s)-\delta/8},$$

where we used that $3s - 2r \gg \sigma \geq \delta$.

Step 1.(b): $K \sim M$. We estimate

$$\begin{aligned} \left\| P_K^u A_{L,\ell}^{-,k} \phi_L^{-,\ell} \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} & \lesssim \|P_K^u A_L^-\|_{L_u^\infty C_v^s} \|\phi_L^-\|_{C_v^{1-r'}} \|\partial_u (H_{M,m}^{+,i} \phi_M^{+,m})\|_{L_u^\infty C_v^s} \\ & \lesssim L^{1-r'-s} M^{1-s} \theta^2 \|P_K^u A_L^-\|_{L_u^\infty C_v^s} \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

For the A_L^- -factor, the condition $K \sim M$ yields that

$$\begin{aligned} \|P_K^u A_L^-\|_{L_u^\infty C_v^s} & \leq \|P_K^u P_{\lesssim L^{1-\delta}}^u A_L^-\|_{L_u^\infty C_v^s} + \|P_K^u P_{\gg L^{1-\delta}}^u A_L^-\|_{L_u^\infty C_v^s} \\ & \lesssim 1\{M \lesssim L^{1-\delta}\} M^{-s} \|P_{\lesssim L^{1-\delta}}^u A_L^-\|_{C_u^s C_v^s} + M^{-r} \|P_{\gg L^{1-\delta}}^u A_L^-\|_{C_u^r C_v^s} \\ & \lesssim (1\{M \lesssim L^{1-\delta}\} M^{-s} + M^{-r}) \|A_L^-\|_{Mod_L^-} \\ & \lesssim (1\{M \lesssim L^{1-\delta}\} M^{-s} + M^{-r}) \theta. \end{aligned}$$

The total pre-factor can then be estimated by

$$\begin{aligned} & L^{1-r'-s} M^{1-s} (1\{M \lesssim L^{1-\delta}\} M^{-s} + M^{-r}) \\ & = 1\{M \lesssim L^{1-\delta}\} L^{1-r'-s} M^{1-2s} + L^{1-r'-s} M^{1-s-r} \\ & \lesssim L^{-\eta} M^{-(1-\delta)^{-1}(r'+s-1)+1-2s} + L^{1-r'-s} M^{1-s-r}. \end{aligned}$$

Since

$$\begin{aligned} -(1-\delta)^{-1}(r'+s-1)+1-2s & \leq -(1+\frac{3\delta}{4})(r'+s-1)+1-2s \\ & = -(r-s) - \frac{3\delta}{4}(r'+s-1) + \eta + 2(1-2s) \\ & \leq -(r-s) - \delta/8, \end{aligned}$$

this yields the desired estimate.

Step 2: The cases (+) and (+−). Using Lemma 4.7, we can assume that

$$\zeta^k(u, v) = E_{L,\ell}^{+,k}(u, v)\phi_L^{+,\ell}(u)$$

for some $L \geq 1$, where E_L^+ satisfies (4.18) or (4.19). In particular, $\|E_L^+\|_{C_u^s C_v^s} \lesssim \theta$. Due to the frequency-support conditions on $E_{L,\ell}^{+,k}$, it follows that $K \sim L$. Then, we decompose

$$P_K^u \left(E_{L,\ell}^{+,k} \phi_L^{+,\ell} \right) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) = E_{L,\ell}^{+,k} P_K^u \phi_L^{+,\ell} H_{M,m}^{+,i} \partial_u \phi_M^{+,m} \quad (5.26)$$

$$+ [P_K^u, E_{L,\ell}^{+,k}] \left(\phi_L^{+,\ell} \right) H_{M,m}^{+,i} \partial_u \phi_M^{+,m} \quad (5.27)$$

$$+ P_K^u \left(E_{L,\ell}^{+,k} \phi_L^{+,\ell} \right) \partial_u H_{M,m}^{+,i} \phi_M^{+,m}. \quad (5.28)$$

We start by treating (5.26), which is the main term. To this end, we estimate

$$\begin{aligned} \left\| E_{L,\ell}^{+,k} P_K^u \phi_L^{+,\ell} H_{M,m}^{+,i} \partial_u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^s} &\lesssim \|E_{L,\ell}^{+,k}\|_{C_u^s C_v^s} \|H_{M,m}^{+,i}\|_{C_u^s C_v^s} \|P_K^u \phi_L^{+,\ell} \partial_u \phi_M^{+,m}\|_{C_u^{r-1}} \\ &\lesssim M^{r-s} L^{-s} \|E_L^+\|_{C_u^s C_v^s} \|H_M^+\|_{C_u^s C_v^s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2 \\ &\lesssim M^{r-s} L^{-s} \theta^3 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

Since $M^{1-\sigma} \leq K \lesssim L$, we have that

$$M^{r-s} L^{-s} \lesssim M^{r-2s+\sigma} L^{-4\eta},$$

which is acceptable. We now turn to the commutator term (5.27). Using Lemma 2.10, we obtain that

$$\begin{aligned} &\left\| [P_K^u, E_{L,\ell}^{+,k}] \left(\phi_L^{+,\ell} \right) H_{M,m}^{+,i} \partial_u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^s} \\ &\lesssim \left\| [P_K^u, E_{L,\ell}^{+,k}] \left(\phi_L^{+,\ell} \right) \right\|_{C_u^{1-r'} C_v^s} \left\| H_{M,m}^{+,i} \partial_u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^s} \\ &\lesssim K^{1-r'-2s+4\eta} M^{r-s} \|E_L^+\|_{C_u^s C_v^s} \|\phi_L^+\|_{C_u^{-4\eta}} \|H_M^+\|_{C_u^s C_v^s} \|\phi_M^+\|_{C_u^s} \\ &\lesssim L^{-4\eta} K^{1-r'-2s+4\eta} M^{r-s} \theta^3 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

Since $K \geq M^{1-\sigma}$, we have that

$$L^{-4\eta} K^{1-r'-2s+4\eta} M^{r-s} \lesssim L^{-4\eta} M^{r-s-(1-\sigma)(1-r-2s+5\eta)} \lesssim L^{-4\eta} M^{1-3s+2\sigma}.$$

Since $1-3s \approx -1/2$ and $r-2s \approx -1/4$, this contribution is acceptable. We now turn to the last term (5.28), which contains a favorable derivative. Using¹² Lemma 4.1, it can be estimated by

$$\begin{aligned} &\left\| P_K^u \left(E_{L,\ell}^{+,k} \phi_L^{+,\ell} \right) \partial_u H_{M,m}^{+,i} \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^s} \\ &\lesssim \left\| P_K^u \left(E_{L,\ell}^{+,k} \phi_L^{+,\ell} \right) \right\|_{C_u^{1-r'} C_v^s} \left\| \partial_u H_{M,m}^{+,i} \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^s} \\ &\lesssim L^{1-r'-s} M^{r-2s} \|E_L^+\|_{C_u^s C_v^s} \|H_M^+\|_{C_u^s C_v^s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s}^2 \end{aligned}$$

¹² While H_M^+ is not necessarily A_M^+ , the same argument applies.

$$\lesssim L^{1-r'-s} M^{r-2s} \theta^3 \|H_M^+\|_{C_u^s C_v^s}.$$

Since $1 - r' - s \approx -1/4$, this term is clearly acceptable.

Step 3: The case (S). We begin with the estimate of the non-resonant portion. To this end, we decompose

$$\begin{aligned} & P_K^u \psi^k \partial_u (H_{M,m}^{+,i} \phi_M^{+,m}) - H_{M,m}^{+,i} (P_K^u \psi^k \odot_\sigma^u \partial_u \phi_M^{+,m}) \\ &= H_{M,m}^{+,i} (P_K^u \psi^k \partial_u \phi_M^{+,m} - P_K^u \psi^k \odot_\sigma^u \partial_u \phi_M^{+,m}) + P_K^u \psi^k \partial_u H_{M,m}^{+,i} \phi_M^{+,m}. \end{aligned} \quad (5.29)$$

The first summand in (5.29) can be estimated using the multiplication estimate (Corollary 2.8) and Lemma 4.5. The second summand in (5.29) can be estimated using the bilinear estimate (Proposition 2.5) and Lemma 4.1.

The estimate of the resonant term $H_{M,m}^{+,i} (P_K^u \psi^k \odot_\sigma^u \partial_u \phi_M^{+,m})$ directly follows from the multiplication estimate (Corollary 2.8) and the inequality (4.13) in Lemma 4.5.

Step 4: The conclusion. If $K \lesssim M$, the combined estimate (5.25) follows directly from the individual estimates and $\max(-(r-s), r-2s+\sigma) < 1-s-r$. In the remaining case $K \gg M$, we have the easier estimate

$$\begin{aligned} \|P_K^u (\zeta^k) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m})\|_{C_u^{r-1} C_v^{1-r'}} &\lesssim K^{r-1} \|P_K^u (\zeta^k) \partial_u (H_{M,m}^{+,i} \phi_M^{+,m})\|_{L_u^\infty C_v^{1-r'}} \\ &\lesssim K^{r-1} \|P_K^u \zeta^k\|_{L_u^\infty C_v^{1-r'}} \|\partial_u (H_{M,m}^{+,i} \phi_M^{+,m})\|_{L_u^\infty C_v^{1-r'}} \\ &\lesssim K^{r-1-s+\eta} M^{r-s} \|\zeta\|_{C_u^{s-\eta} C_v^{1-r'}} \|H_M^+\|_{C_u^s C_v^s} \|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \\ &\lesssim K^{2r-2+1-2s+\eta} \mathcal{G}(\zeta) \theta^2 \|H_M^+\|_{C_u^s C_v^s}. \end{aligned}$$

Since $1 - 2s \leq \eta$, this yields the desired estimate. \square

Corollary 5.8. *Let Hypothesis 3.8 be satisfied, let $1 \leq a, b, m \leq D$ and let K, L, M be frequency-scales satisfying $K \sim M$. Furthermore, let $\zeta_1: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ and let ζ_2 be of type $(-)$, $(+)$, $(+-)$, or (S) . Then, it holds that*

$$\|P_L^v (\zeta_1^a P_K^u \zeta_2^b) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \lesssim (M^{-(r-s)-\delta/8} + L^{1-s-r+2\eta} M^{1-s-r}) \|\zeta_1\|_{C_u^s C_v^s} \mathcal{G}(\zeta_2) \theta^2.$$

Proof. When applying Lemma 5.6 in the following argument, we always take $H_M^+ = \text{Id}$. We distinguish two different cases according to the type of ζ_2 .

Case 1: ζ_2 is of type $(-)$. We then decompose $\zeta_1^a = P_{\ll L}^v \zeta_1^a + P_{\gtrsim L}^v \zeta_1^a$. To estimate the contribution of the low-frequency term, we note that

$$\begin{aligned} \|P_L^v (P_{\ll L}^v \zeta_1^a P_K^u \zeta_2^b) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} &= \|P_L^v (P_{\ll L}^v \zeta_1^a P_K^u \tilde{P}_L^v \zeta_2^b \partial_u \phi_M^{+,m})\|_{C_u^{r-1} C_v^{1-r'}} \\ &\leq \|P_{\ll L}^v \zeta_1^a P_K^u \tilde{P}_L^v \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \\ &\leq \|\zeta_1^a\|_{C_u^s C_v^s} \|P_K^u \tilde{P}_L^v \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \\ &= \|\zeta_1^a\|_{C_u^s C_v^s} \|\tilde{P}_L^v (P_K^u \zeta_2^b \partial_u \phi_M^{+,m})\|_{C_u^{r-1} C_v^{1-r'}}. \end{aligned}$$

The desired estimate then follows from the $(-)$ -estimate in Lemma 5.6. For the high-frequency term in ζ_1 , we estimate

$$\|P_L^v (P_{\gtrsim L}^v \zeta_1^a P_K^u \zeta_2^b) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} = \|P_L^v (P_{\gtrsim L}^v \zeta_1^a P_K^u \zeta_2^b \partial_u \phi_M^{+,m})\|_{C_u^{r-1} C_v^{1-r'}}$$

$$\begin{aligned}
&\leq \|P_{\gtrsim L}^v \zeta_1^a P_K^u \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \\
&\lesssim \|P_{\gtrsim L}^v \zeta_1^a\|_{C_u^{1-r'} C_v^{1-r'}} \|P_K^u \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \\
&\lesssim L^{1-r'-s} \|\zeta_1^a\|_{C_u^s C_v^s} \|P_K^u \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}}.
\end{aligned}$$

The desired estimate then follows from (5.25).

Case 2: ζ_2 is of type (+), (+−), or (S). In this case, we estimate

$$\|P_L^v (\zeta_1^a P_K^u \zeta_2^b) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \lesssim L^{1-r'-s} \|\zeta_1^a\|_{C_u^s C_v^s} \|P_K^u \zeta_2^b \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^s}.$$

Thus, the desired estimate follows from the (+), (+−), or (S)-estimate in Lemma 5.6. \square

6. The Full Nonlinearity

In the previous section, we analyzed both the null form $\partial_u \phi^i \partial_v \phi^j$ and the product $\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u \phi^i$. In this section, we treat the full $\mathbb{I}_{ij}^{\diamond,k}(\phi) \partial_u \phi^i \partial_v \phi^j$. For the most part, the desired estimates follow directly by combining estimates from Sect. 5. The primary exception is the term

$$\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}), \quad (6.1)$$

which was only briefly considered in Proposition 5.1. As was mentioned previously, the reason for not considering (6.1) before is that the case-analysis is more efficient when $\mathbb{I}_{ij}^{\diamond,a}(\phi)$ is already included. The main result of this section is the following theorem.

Theorem 6.1 (Nonlinearity). *Let Hypothesis 3.8 be satisfied and let $1 \leq a \leq D$. Then, it holds that*

$$\begin{aligned}
\mathbb{I}_{ij}^a(\phi) \partial_u \phi^i \partial_v \phi^j &= - \sum_{M \sim_\delta N} G_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} + \sum_M F_{M,m}^{+,a} \mathbb{E}_\sigma^u \partial_u \phi_M^{+,m} \\
&\quad + \sum_N F_{N,n}^{-,a} \mathbb{E}_\sigma^v \partial_v \phi_N^{-,n} + \mathcal{E}^a,
\end{aligned}$$

where $G_{M,N,mn}^a$, $F_{M,m}^{+,a}$, and $F_{N,n}^{-,a}$ are as in Definition 3.12 and the remainder term \mathcal{E} satisfies

$$\|\mathcal{E}^a\|_{C_u^{r-1} C_v^{r-1}} \lesssim \theta^2.$$

6.1. The main term. We start by analyzing the main term (6.1). By symmetry in the u and v -variables, we can assume that $N \leq M$. We further split the analysis in the two cases $N \leq M^{1-\delta}$ and $M^{1-\delta} < N \leq M$.

6.1.1. *The case $N \leq M^{1-\delta}$* This case is relatively simple, as it will only contribute to the modulated version of $\phi_M^{+,m}$.

Proposition 6.2 (Main term for $N \leq M^{1-\delta}$). *Let Hypothesis 3.8 be satisfied, let $1 \leq a \leq D$, and let M and N be frequency scales satisfying $N \leq M^{1-\delta}$. Then, it holds that*

$$\begin{aligned} & \left\| \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right. \\ & \quad \left. - \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \partial_u (\phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim (MN)^{-\eta} \theta^4. \end{aligned} \quad (6.2)$$

Proof. We start the proof with two steps which reduce (6.2) to a frequency-localized version which does not contain para-product operators. Using Corollary 4.2, it holds that

$$\begin{aligned} & \left\| \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right. \\ & \quad \left. - \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \partial_u (\phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M^{r-2s+\sigma} \left\| \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^s C_v^{r-1}} \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\phi_M^{+,m}\|_{C_u^s} \\ & \lesssim M^{r-2s+\sigma} N^{r-s} \theta^4. \end{aligned}$$

Since $N \leq M$ and $3s - 2r \gg \sigma$, this is an acceptable contribution to (6.2). Using Lemma 2.14, it holds that

$$\begin{aligned} & \left\| \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) - P_{\leq M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M^{-(1-\sigma)s} \left\| \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^s C_v^{r-1}} \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \\ & \lesssim M^{-(1-\sigma)s} N^{r-s} M^{r-s} \theta^4 \\ & \lesssim (MN)^{-\eta} \theta^4, \end{aligned}$$

where we used $3s - 2r \gg \sigma$. In order to prove (6.2), it therefore remains to prove that

$$\left\| P_{>M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim (MN)^{-\eta} \theta^4.$$

Using Lemma 2.10, it holds that

$$P_{>M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) = P_{>M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \right) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \quad (6.3)$$

$$+ \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v P_{>M^{1-\sigma}}^u (A_{N,n}^{-,j} \phi_N^{-,n}) \quad (6.4)$$

$$+ \mathcal{E}_{M,N,i}^a, \quad (6.5)$$

where the error term satisfies

$$\|\mathcal{E}_{M,N,i}^a\|_{C_u^{1-r'} C_v^{r-1}} \lesssim M^{(1-\sigma)(1-r'-2s)} N^{r-s} \theta^2.$$

We now estimate the contributions of (6.3), (6.4), and (6.5) separately.

Step 1: Contribution of (6.3). Using the bilinear estimate (Proposition 2.5) and the product estimate (Lemma 5.6), it holds that

$$\left\| P_{>M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \right) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}}$$

$$\begin{aligned} &\lesssim \left\| P_{>M^{1-\sigma}}^u \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) \right) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} \left\| \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{1-r'} C_v^{r-1}} \\ &\lesssim M^{1-s-r+\eta} N^{r-s} \theta^4. \end{aligned}$$

Since $N \leq M^{1-\delta}$, we have that

$$M^{1-s-r+\eta} N^{r-s} \lesssim M^{1-s-r+\eta+(1-\delta)(r-s)} = M^{1-2s+\eta-\delta(r-s)} \lesssim M^{-\delta/6}, \quad (6.6)$$

which is acceptable.

Step 2: Contribution of (6.4). We first note that

$$P_{>M^{1-\sigma}}^u (A_{N,n}^{-,j} \phi_N^{-,n}) = (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j}) \phi_N^{-,n}.$$

Then, we recall that $N^{1-\delta} \ll N \leq M^{1-\delta} \ll M$. Using the definition of Mod_N^- , it follows that

$$\|A_{N,n}^{-,j}\|_{C_u^s C_v^s} + \|\tilde{P}_M^u A_{N,n}^{-,j}\|_{C_u^r C_v^s} \lesssim \theta.$$

Using the multiplication estimate (Corollary 2.8), we have that

$$\begin{aligned} &\left\| \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}}. \end{aligned}$$

We now further separate the resonant and non-resonant terms in the u -variable. Using the bilinear estimate (Proposition 2.5.(iv)), the resonant term can be estimated by

$$\begin{aligned} &\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \odot^u \partial_v (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{-r'} C_v^s} \left\| \partial_v (\tilde{P}_M^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^r C_v^s} \\ &\lesssim M^{1-s-r'} N^{r-s} \|A_M^+\|_{C_u^s C_v^s} \|\phi^+\|_{C_u^s} \|\tilde{P}_M^u A_N^-\|_{C_u^r C_v^s} \|\phi^-\|_{C_v^s} \\ &\lesssim M^{1-s-r'} N^{r-s} \theta^4. \end{aligned}$$

Using (6.6), this term is acceptable. Using Proposition 2.5.(iii), the non-resonant term can be estimated by

$$\begin{aligned} &\left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \oslash^u \partial_v (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^s} \left\| \partial_v (P_{>M^{1-\sigma}}^u A_{N,n}^{-,j} \phi_N^{-,n}) \right\|_{C_u^\eta C_v^{r-1}} \\ &\lesssim M^{r-s} N^{r-s} \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\phi_M^{+,m}\|_{C_u^s} \|P_{>M^{1-\sigma}}^u A_{N,n}^{-,j}\|_{C_u^\eta C_v^s} \|\phi_N^{-,n}\|_{C_v^s} \\ &\lesssim M^{r-s} M^{(1-\sigma)(\eta-s)} N^{r-s} \theta^4. \end{aligned}$$

Since $N \leq M$ and $3s - 2r \gg \sigma$, this is acceptable.

Step 3: Contribution of (6.5). Using the bilinear estimate (Proposition 2.5), we have that

$$\begin{aligned} \left\| \mathcal{E}_{M,N,i}^a \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} &\lesssim \left\| \mathcal{E}_{M,N,i}^a \right\|_{C_u^{1-r'} C_v^{r-1}} \left\| \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{1-r'}} \\ &\lesssim M^{(1-\sigma)(1-r'-2s)} M^{r-s} N^{r-s} \theta^4. \end{aligned}$$

Since $N \leq M$ and

$$(1-\sigma)(1-r'-2s) + r-s + r-s \approx -1/4,$$

this term is acceptable.

□

6.1.2. The case $M \sim_\delta N$ This case is slightly harder, since it contributes to the bilinear term containing $\phi_M^{+,m} \phi_N^{-,n}$, the modulated version of $\phi_M^{+,m}$, and the modulated version of $\phi_N^{-,n}$.

Proposition 6.3 (Main term for $M \sim_\delta N$). *Let Hypothesis 3.8 be satisfied, let $1 \leq i, j \leq D$, and let M, N be two frequency-scales satisfying $M \sim_\delta N$. For all $1 \leq a, m, n \leq D$, define*

$$\zeta_{M,N,mn}^a := \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}. \quad (6.7)$$

Then, it holds that

$$\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) = -G_{M,N,mn}^a (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \quad (6.8)$$

$$+ (\zeta_{M,N,mn}^a \odot^v \partial_v \phi_N^{-,n}) \otimes_\sigma^u \partial_u \phi_M^{+,m} \quad (6.9)$$

$$+ (\zeta_{M,N,mn}^a \odot^u \partial_u \phi_M^{+,m}) \otimes_\sigma^v \partial_v \phi_N^{-,n} \quad (6.10)$$

$$+ \mathcal{E}_{M,N}^a, \quad (6.11)$$

where $G_{M,N,mn}^a$ is as in Definition 3.12 and the error term $\mathcal{E}_{M,N}^a$ satisfies

$$\|\mathcal{E}_{M,N}^a\|_{C_u^{r-1} C_v^{r-1}} \lesssim (MN)^{-\eta} \theta^4.$$

Before starting with the proof of Proposition 6.3, we require the following auxiliary lemma. In Lemma 6.4, the function $\zeta_{M,N,mn}^a$ from (6.7) is replaced by a general function $\zeta \in C_u^s C_v^s$. In the spirit of Definition 2.4, we define the para-product operators $\otimes_\sigma^{u,v}$ and $\odot^{u,v}$ by

$$f \otimes_\sigma^{u,v} g := \sum_{\substack{M_1, N_1: \\ M_1 \leq N_1^{1-\sigma}}} \sum_{\substack{M_2, N_2: \\ M_2 \leq N_2^{1-\sigma}}} P_{M_1}^u P_{M_2}^v f P_{N_1}^u P_{N_2}^v g,$$

$$f \odot^{u,v} g := \sum_{\substack{M_1, N_1: \\ M_1 \sim N_1}} \sum_{\substack{M_2, N_2: \\ M_2 \sim N_2}} P_{M_1}^u P_{M_2}^v f P_{N_1}^u P_{N_2}^v g.$$

Lemma 6.4. *Let Hypothesis 3.8 be satisfied. Let $1 \leq m, n \leq D$, let M, N be frequency-scales satisfying $M \sim_\delta N$, and let $\zeta \in C_u^s C_v^s$. Then, we can write*

$$\zeta \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} = \zeta \otimes_\sigma^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \quad (6.12)$$

$$+ (\zeta \odot^v \partial_v \phi_N^{-,n}) \otimes_\sigma^u \partial_u \phi_M^{+,m} \quad (6.13)$$

$$+ (\zeta \odot^u \partial_u \phi_M^{+,m}) \otimes_\sigma^v \partial_v \phi_N^{-,n} \quad (6.14)$$

$$+ \zeta \odot^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \quad (6.15)$$

$$+ \mathcal{E}_{M,N}^{m,n}, \quad (6.16)$$

where the error term $\mathcal{E}_{M,N}^{m,n}$ satisfies

$$\|\mathcal{E}_{M,N}^{m,n}\|_{C_u^{r-1} C_v^{r-1}} \lesssim (MN)^{-\eta} \theta^2 \|\zeta\|_{C_u^s C_v^s}.$$

Proof of Lemma 6.4: Using a Littlewood-Paley decomposition, we write

$$\zeta = \sum_{K,L} P_K^u P_L^v \zeta.$$

We now split all possible frequency-scales K and L into seven regions as follows:

$$2^{\mathbb{N}_0} \times 2^{\mathbb{N}_0} = \{(K, L): K \leq M^{1-\sigma}, L \leq N^{1-\sigma}\} \quad (6.17)$$

$$\dot{\cup} \{(K, L): K \leq M^{1-\sigma}, L \sim N\} \quad (6.18)$$

$$\dot{\cup} \{(K, L): K \sim M, L \leq N^{1-\sigma}\} \quad (6.19)$$

$$\dot{\cup} \{(K, L): K \sim M, L \sim N\} \quad (6.20)$$

$$\dot{\cup} \{(K, L): K \sim M, N^{1-\sigma} < L \not\sim N\} \quad (6.21)$$

$$\dot{\cup} \{(K, L): M^{1-\sigma} < K \not\sim M, L \sim N\} \quad (6.22)$$

$$\dot{\cup} \{(K, L): K > M^{1-\sigma} \text{ or } L > N^{1-\sigma}, K \not\sim M, L \not\sim N\}. \quad (6.23)$$

The contributions of the regions (6.17)-(6.20) will lead to the main terms (6.12)-(6.15). The remaining contributions of (6.21), (6.22), and (6.23) will contribute exclusively towards the error term $\mathcal{E}_{M,N}^{m,n}$. We now treat the individual contributions separately.

Contribution of (6.17): The contribution of this frequency region is given by

$$\sum_{K \leq M^{1-\sigma}} \sum_{L \leq N^{1-\sigma}} P_K^u P_L^v \zeta \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} = (P_{\leq M^{1-\sigma}}^u P_{\leq N^{1-\sigma}}^v \zeta) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}.$$

Using a variant of Lemma 2.14, it holds that

$$\begin{aligned} & \left\| (P_{\leq M^{1-\sigma}}^u P_{\leq N^{1-\sigma}}^v \zeta) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} - \zeta \otimes_{\sigma}^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \min(M, N)^{-(1-\sigma)s} M^{r-s} N^{r-s} \|\zeta\|_{C_u^s C_v^s} \theta^2. \end{aligned}$$

Since $M \sim_{\sigma} N$ and $3s - 2r \gg \sigma$, this is acceptable.

Contributions of (6.18), (6.19), and (6.20): The argument is similar as for the contribution of (6.17) and only relies on Lemma 2.14.

Contribution of (6.21): Using Proposition 2.5.(iii), we have that

$$\begin{aligned} & \sum_{\substack{K,L: \\ K \sim M, \\ N^{1-\sigma} < L \not\sim N}} \left\| P_K^u P_L^v \zeta \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \sum_{\substack{K,L: \\ K \sim M, \\ N^{1-\sigma} < L \not\sim N}} \left\| P_K^u P_L^v \zeta \right\|_{C_u^{\eta} C_v^{\eta}} \left\| \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} \right\|_{C_u^{\eta} C_v^{r-1}} \\ & \lesssim \sum_{\substack{K,L: \\ K \sim M, \\ N^{1-\sigma} < L \not\sim N}} K^{\eta-s} L^{\eta-s} M^{1+\eta-s} N^{r-s} \theta^2 \|\zeta\|_{C_u^s C_v^s} \\ & \lesssim M^{1+2\eta-2s} N^{r-s-(1-\sigma)(\eta-s)} \theta^2 \|\zeta\|_{C_u^s C_v^s}. \end{aligned}$$

Since $M \sim_\delta N$ and

$$1 + 2\eta - 2s + r - s - (1 - \sigma)(\eta - s) \approx -1/4,$$

this is acceptable.

Contribution of (6.22): Using the symmetry in the u and v -variables, this follows from the same argument as for (6.21).

Contribution of (6.23): By symmetry, it suffices to estimate the contribution for $K > M^{1-\sigma}$. Using Proposition 2.5.(iii), it holds that

$$\begin{aligned} & \sum_{\substack{K, L: \\ M^{1-\sigma} < K \not\sim M, \\ L \not\sim N}} \|P_K^u P_L^v \zeta \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \sum_{\substack{K, L: \\ M^{1-\sigma} < K \not\sim M, \\ L \not\sim N}} \|P_K^u P_L^v \zeta\|_{C_u^\eta C_v^{1-r'}} \|\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \sum_{\substack{K, L: \\ M^{1-\sigma} < K \not\sim M, \\ L \not\sim N}} K^{\eta-s} L^{1-r'-s} M^{r-s} N^{r-s} \theta^2 \|\zeta\|_{C_u^s C_v^s} \\ & \lesssim M^{(1-\sigma)(\eta-s)+r-s} N^{r-s} \theta^2 \|\zeta\|_{C_u^s C_v^s}. \end{aligned}$$

Since $M \sim_\delta N$ and $3s - 2r \gg \sigma$, this term is acceptable. \square

Equipped with Lemma 6.4, we can now prove Proposition 6.3.

Proof of Proposition 6.3.: We separate the proof into several steps.

Step 1: Removing terms with favorable derivatives. Using the multiplication estimate (Corollary 2.8) and Proposition 5.1, it holds that

$$\begin{aligned} & \|\Pi_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) - \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \|\Pi_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) - A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim (MN)^{-\eta} \theta^4. \end{aligned}$$

Thus, we have reduced the main term to

$$\Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} = \zeta_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n},$$

where $\zeta_{M,N,mn}^a$ is as in (6.7).

Step 2: Applying Lemma 6.4. Using Lemma 6.4, it holds that

$$\Pi_{ij}^{\diamond,a}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) = \zeta_{M,N,mn}^a \ll_{\sigma}^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \quad (6.24)$$

$$+ (\zeta_{M,N,mn}^a \sim^v \partial_v \phi_N^{-,n}) \ll_{\sigma}^u \partial_u \phi_M^{+,m} \quad (6.25)$$

$$+ (\zeta_{M,N,mn}^a \sim^u \partial_u \phi_M^{+,m}) \ll_{\sigma}^v \partial_v \phi_N^{-,n} \quad (6.26)$$

$$\begin{aligned} & + \zeta_{M,N,mn}^a \sim^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}) \\ & + \tilde{\mathcal{E}}_{M,N}^a, \end{aligned} \quad (6.27)$$

where the error term satisfies

$$\|\tilde{\mathcal{E}}_{M,N}^a\|_{C_u^{r-1}C_v^{r-1}} \lesssim (MN)^{-\eta}\theta^2 \|\zeta_{M,N,mn}^a\|_{C_u^sC_v^s} \lesssim (MN)^{-\eta}\theta^4.$$

Using Lemma 2.14 and Definition 3.12, the first summand (6.24) leads to (6.8) and an acceptable error term. The second and third summands (6.25) and (6.26) coincide with (6.9) and (6.10), respectively. As a result, it only remains to prove that

$$\|\zeta_{M,N,mn}^a \odot^{u,v} (\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n})\|_{C_u^{r-1}C_v^{r-1}} \lesssim (MN)^{-\eta}\theta^4. \quad (6.28)$$

The double-resonance estimate (6.28) requires much more detailed information than only $\zeta_{M,N,mn}^a \in C_u^sC_v^s$, which explains why this term was not estimated in Lemma 6.4.

Step 3: The double-resonance estimate (6.28). Using a variant of Lemma 2.14, it suffices to estimate

$$\sum_{\substack{K,L: \\ K \sim M, L \sim N}} P_K^u P_L^v \zeta_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}.$$

In the following, we fix frequency scales $K \sim M$ and $L \sim N$. Using (2.44) in Lemma 2.10, we have that

$$P_K^u P_L^v \zeta_{M,N,mn}^a = P_K^u P_L^v (\Pi_{ij}^{\diamond,a}(\phi)) A_{M,m}^{+,i} A_{N,n}^{-,j} \quad (6.29)$$

$$+ P_K^u (\Pi_{ij}^{\diamond,a}(\phi)) P_L^v (A_{M,m}^{+,i} A_{N,n}^{-,j}) \quad (6.30)$$

$$+ P_L^v (\Pi_{ij}^{\diamond,a}(\phi)) P_K^u (A_{M,m}^{+,i} A_{N,n}^{-,j}) \quad (6.31)$$

$$+ \Pi_{ij}^{\diamond,a}(\phi) P_K^u P_L^v (A_{M,m}^{+,i} A_{N,n}^{-,j}) \quad (6.32)$$

$$+ \tilde{\zeta}_{M,N,mn}^a, \quad (6.33)$$

where $\tilde{\zeta}_{M,N,mn}^a$ satisfies

$$\begin{aligned} \|\tilde{\zeta}_{M,N,mn}^a\|_{C_u^sC_v^s} &\lesssim \min(K, L)^{-s} \|\Pi_{ij}^{\diamond,a}(\phi)\|_{C_u^sC_v^s} \|A_{M,m}^{+,i}\|_{C_u^sC_v^s} \|A_{N,n}^{-,j}\|_{C_u^sC_v^s} \\ &\lesssim \min(M, N)^{-s}\theta^2. \end{aligned}$$

The contribution of the error term $\tilde{\zeta}_{M,N,mn}^a$ can be estimated by

$$\begin{aligned} \|\tilde{\zeta}_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1}C_v^{r-1}} &\lesssim \|\tilde{\zeta}_{M,N,mn}^a\|_{C_u^sC_v^s} \|\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1}C_v^{r-1}} \\ &\lesssim \min(M, N)^{-s} M^{r-s} N^{r-s} \theta^4. \end{aligned}$$

Since $M \sim_\delta N$ and $3s - 2r \gg \delta$, this contribution is acceptable. Thus, it remains to treat the contributions of (6.29), (6.30), (6.31), and (6.32), which will be handled separately.

Step 3.(a): Contribution of (6.29). Due to the multiplication estimate (Corollary 2.8), it suffices to prove that

$$\|P_K^u P_L^v (\Pi_{ij}^{\diamond,a}(\phi)) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1}C_v^{r-1}} \lesssim (MN)^{-\eta}\theta^2. \quad (6.34)$$

Using the symmetry in the u and v -variables, we can assume that $M \geq N$. We now break the symmetry of the term in (6.34) by decomposing

$$P_K^u P_L^v (\Pi_{ij}^{\diamond,a}(\phi)) = P_L^v (\partial_k \Pi_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) + P_L^v (P_K^u \Pi_{ij}^{\diamond,a}(\phi) - \partial_k \Pi_{ij}^{\diamond,a}(\phi) P_K^u \phi^k). \quad (6.35)$$

Using Bony's parilinearization (Lemma 2.9), the second summand in (6.35) is controlled by

$$\begin{aligned} & \|P_L^v(P_K^u \mathbb{I}_{ij}^{\diamond,a}(\phi) - \partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k)\|_{C_u^s C_v^{1-r'}} \\ & \lesssim L^{1-r'-s} \|P_L^v(P_K^u \mathbb{I}_{ij}^{\diamond,a}(\phi) - \partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k)\|_{C_u^s C_v^s} \\ & \lesssim L^{1-r'-s} K^{-s}. \end{aligned}$$

By using the bilinear estimate (Proposition 2.5) and Proposition 5.1, this yields an acceptable contribution to (6.34). Thus, we have further reduced (6.34) to the estimate

$$\|P_L^v(\partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \lesssim (MN)^{-\eta} \theta^2. \quad (6.36)$$

Using the bilinear estimate (Proposition 2.5), it holds that

$$\begin{aligned} & \|P_L^v(\partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \|P_L^v(\partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \|\partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \\ & \lesssim N^{r-s} \|P_L^v(\partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \theta. \end{aligned}$$

Using $K \sim M$, $L \sim N$, and Corollary 5.8, we have that

$$\begin{aligned} & N^{r-s} \|P_L^v(\partial_k \mathbb{I}_{ij}^{\diamond,a}(\phi) P_K^u \phi^k) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \theta \\ & \lesssim N^{r-s} (M^{-(r-s)-\delta/8} + N^{1-s-r+2\eta} M^{1-s-r}) \theta^3 \\ & \lesssim ((NM^{-1})^{r-s} M^{-\delta/8} + N^{1-2s+2\eta} M^{1-s-r}) \theta^3. \end{aligned}$$

Using our assumption $M \geq N$, this term is acceptable.

Step 3.(b): Contribution of (6.30). Using the multiplication estimate (Corollary 2.8) and the bilinear estimate (Proposition 2.5), we have that

$$\begin{aligned} & \|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) P_L^v(A_{M,m}^{+,i} A_{N,n}^{-,j}) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \|P_L^v(A_{M,m}^{+,i} A_{N,n}^{-,j})\|_{C_u^{1-r'} C_v^{1-r'}} \|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim L^{1-r'-s} \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|A_{N,n}^{-,j}\|_{C_u^s C_v^s} \|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \|\phi_N^{-,n}\|_{C_v^{r-1}} \\ & \lesssim L^{1-r'-s} N^{r-s} \|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \theta^3. \end{aligned} \quad (6.37)$$

Using Lemma 5.6 with $H = \text{Id}$ and Bony's parilinearization, it holds that

$$\|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_u \phi_M^{+,m}\|_{C_u^{r-1} C_v^{1-r'}} \lesssim M^{1-s-r} \theta. \quad (6.38)$$

By combining (6.37) and (6.38), it follows that

$$\|P_K^u(\mathbb{I}_{ij}^{\diamond,a}(\phi)) P_L^v(A_{M,m}^{+,i} A_{N,n}^{-,j}) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \lesssim L^{1-r'-s} N^{r-s} M^{1-s-r} \theta^4.$$

Since $L \sim N$, $M \sim_\delta N$, and

$$1 - r' - s + r - s + 1 - s - r = 2 - 3s - r + \eta \approx -1/4,$$

this is acceptable.

Step 3.(c): Contribution of (6.31). By symmetry in the u and v -variable, this follows from the same argument as in Step 3.(b).

Step 3.(d): Contribution of (6.32). By symmetry, we may assume that $M \geq N$. Since $A_{M,m}^{+,i}$ only has u -frequencies bounded by $M^{1-\sigma}$, we can replace $A_{N,n}^{-,j}$ by $P_{\gg N^{1-\delta}}^u A_{N,n}^{-,j}$. It follows that

$$\begin{aligned} & \|\Pi_{ij}^{\diamond,a}(\phi) P_K^u P_L^v (A_{M,m}^{+,i} P_{\gg N^{1-\delta}}^u A_{N,n}^{-,j}) \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \|\Pi_{ij}^{\diamond,a}(\phi)\|_{L_{u,v}^\infty} \|P_K^u P_L^v (A_{M,m}^{+,i} P_{\gg N^{1-\delta}}^u A_{N,n}^{-,j})\|_{L_{u,v}^\infty} \|\partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}\|_{L_{u,v}^\infty} \\ & \lesssim K^{-s} L^{-s} M^{1-s} N^{1-s} N^{-(1-\delta)(r-s)} \|A_M^+\|_{C_u^s C_v^s} \|P_{\gg N^{1-\delta}}^u A_N^-\|_{C_u^r C_v^r} \|\phi^+\|_{C_u^s} \|\phi^-\|_{C_v^s} \\ & \lesssim M^{1-2s} N^{1-2s} N^{-(1-\delta)(r-s)} \theta^4. \end{aligned}$$

Since $M \sim_\delta N$ and $2(1-2s) - (1-\delta)(r-s) \approx -1/4$, this term is acceptable. \square

6.2. The para-controlled and resonance terms. We now control the terms $\Pi_{ij}^{\diamond,k}(\phi) \partial_u \zeta_1^i \partial_v \zeta_2^j$, where the tuple (ζ_1, ζ_2) falls into the \otimes_σ^u and $\neg \odot_\sigma^v$ -cases in Fig. 2.

Proposition 6.5 (The para-controlled term) *Let Hypothesis 3.8 be satisfied, let $1 \leq k \leq D$, and let M be a frequency scale. Furthermore, let ζ be of type $(+)$, $(+-)$, or (S) . Then, it holds that*

$$\begin{aligned} & \left\| \Pi_{ij}^{\diamond,k}(\phi) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v \zeta^j - (\Pi_{ij}^{\diamond,k}(\phi) A_{M,m}^{+,i} \partial_v \zeta^j) \otimes_\sigma^u \partial_u (\phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim M^{-\eta} \mathcal{G}(\zeta) \theta^3. \end{aligned} \quad (6.39)$$

We also have the frequency-localized variant

$$\left\| P_K^u (\Pi_{ij}^{\diamond,k}(\phi) \partial_v \zeta^j) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim K^{-\eta} \mathcal{G}(\zeta) \theta^3. \quad (6.40)$$

for all frequency-scales K satisfying $K \gtrsim M^{1-\sigma}$.

Proof. Due to Lemma 2.14, Corollary 4.2, and elementary estimates, it suffices to prove (6.40). To this end, we split we split $P_K^u (\Pi_{ij}^{\diamond,k}(\phi) \partial_v \zeta^j)$ into a low \times high, a high \times low, and an error term. To be precise, we write

$$P_K^u (\Pi_{ij}^{\diamond,k}(\phi) \partial_v \zeta^j) =: P_{\ll K}^u (\Pi_{ij}^{\diamond,k}(\phi)) \partial_v P_K^u \zeta^j + P_K^u (\Pi_{ij}^{\diamond,k}(\phi)) \partial_v P_{\ll K}^u \zeta^j + \mathcal{E}_{K,i}^k, \quad (6.41)$$

which serves as a definition of the error term $\mathcal{E}_{K,i}^k$. We now split the argument into three steps corresponding to the three summands in (6.41).

Step 1: The contribution of the low \times high-term. Using the multiplication estimate (Corollary 2.8) and the composition estimate (Lemma 2.9), we have that

$$\begin{aligned} & \left\| P_{\ll K}^u (\Pi_{ij}^{\diamond,k}(\phi)) \partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v P_K^u \zeta^j \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \|\Pi_{ij}^{\diamond,k}(\phi)\|_{C_u^s C_v^s} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v P_K^u \zeta^j\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim (1 + \|\phi\|_{C_u^s C_v^s})^{10} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m}) \partial_v P_K^u \zeta^j\|_{C_u^{r-1} C_v^{r-1}} \end{aligned}$$

$$\lesssim \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\partial_v P_K^u \zeta^j\|_{C_u^{r-1}C_v^{r-1}}.$$

It remains to prove that

$$\begin{aligned} & \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\partial_v P_K^u \zeta^j\|_{C_u^{r-1}C_v^{r-1}} \\ & \lesssim K^{-\eta} \mathcal{G}(\zeta) \theta^3. \end{aligned}$$

If $M^{1-\sigma} \lesssim K \lesssim M$, this directly follows from Proposition 5.3. If $K \gg M$, this follows easily from Proposition 2.5.(iii). Indeed,

$$\begin{aligned} \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\partial_v P_K^u \zeta^j\|_{C_u^{r-1}C_v^{r-1}} & \lesssim \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^\eta C_v^\delta} \|\partial_v P_K^u \zeta^j\|_{C_u^{r-1}C_v^{r-1}} \\ & \lesssim M^{1-s+\eta} K^{r-1-(1-r')} \|A_M^+\|_{C_u^\delta C_v^\delta} \|\phi^+\|_{C_u^\delta} \|P_K^u \zeta\|_{C_u^{1-r'}C_v^{r-1}} \\ & \lesssim K^{2r-1-s+2\eta} \theta^2 \|P_K^u \zeta\|_{C_u^{1-r'}C_v^{r-1}}. \end{aligned} \quad (6.42)$$

By inserting the types (+), (+−), or (S), it is easy to see that

$$\|P_K^u \zeta\|_{C_u^{1-r'}C_v^{r-1}} \lesssim K^{2\delta} \mathcal{G}(\zeta) \theta.$$

After inserting this back into (6.42) and using $3/4 - r \gg \delta$, this completes the first step.

Step 2: The contribution of the high \times low-term. Using the standard bilinear estimate (Proposition 2.5), we have that

$$\begin{aligned} & \|P_K^u(\Pi_{ij}^{\diamond,k}(\phi))\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\partial_v P_K^u \zeta^j\|_{C_u^{r-1}C_v^{r-1}} \\ & \lesssim \|P_K^u(\Pi_{ij}^{\diamond,k}(\phi))\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r-1}C_v^{1-r'}} \|\partial_v P_K^u \zeta^j\|_{C_u^{1-r'}C_v^{r-1}}. \end{aligned} \quad (6.43)$$

Similar as in Step 1, it is easy to see for ζ of type (+), (+−), or (S) that

$$\|\partial_v P_K^u \zeta^j\|_{C_u^{1-r'}C_v^{r-1}} \lesssim K^{2\delta} \mathcal{G}(\zeta) \theta. \quad (6.44)$$

The 2δ -loss in (6.44) will easily be absorbed through our estimate of the first factor in (6.44), which we now present. Inspired by Bony's parilinearization formula, we decompose

$$P_K^u(\Pi_{ij}^{\diamond,k}(\phi)) = \left((\partial_\ell \Pi_{ij}^{\diamond,k})(\phi) P_K^u \phi^\ell \right) + \left(P_K^u(\Pi_{ij}^{\diamond,k}(\phi)) - (\partial_\ell \Pi_{ij}^{\diamond,k})(\phi) P_K^u \phi^\ell \right). \quad (6.45)$$

The contribution of the first term in (6.45) is estimated using the multiplication estimate (Corollary 2.8) and the product estimate (Lemma 5.6), which yield

$$\begin{aligned} & \|(\partial_\ell \Pi_{ij}^{\diamond,k})(\phi) P_K^u \phi^\ell \partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r-1}C_v^{1-r'}} \\ & \lesssim \|(\partial_\ell \Pi_{ij}^{\diamond,k})(\phi)\|_{C_u^\delta C_v^\delta} \|P_K^u \phi^\ell \partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r-1}C_v^{1-r'}} \\ & \lesssim (1\{K \lesssim M\} M^{1-s-r} + 1\{K \gg M\} K^{2(r-1+\eta)}) \theta^3. \end{aligned} \quad (6.46)$$

Since $1 - s - r \approx -1/4$ and $2(r - 1 + \eta) \approx -1/2$, the gain in (6.46) easily beats the 2δ -loss in (6.44). The contribution of the second term in (6.45) is estimated using Bony's parilinearization estimate (Lemma 2.9), which yields

$$\begin{aligned} & \left\| \left(P_K^u(\Pi_{ij}^{\diamond,k}(\phi)) - (\partial_\ell \Pi_{ij}^{\diamond,k})(\phi) P_K^u \phi^\ell \right) \partial_u(A_{M,m}^{+,i}\phi_M^{+,m}) \right\|_{C_u^{r-1}C_v^{1-r'}} \\ & \lesssim \|P_K^u(\Pi_{ij}^{\diamond,k}(\phi)) - (\partial_\ell \Pi_{ij}^{\diamond,k})(\phi) P_K^u \phi^\ell\|_{C_u^\delta C_v^\delta} \|\partial_u(A_{M,m}^{+,i}\phi_M^{+,m})\|_{C_u^{r-1}C_v^\delta} \\ & \lesssim K^{-s} M^{r-s} \theta^3. \end{aligned} \quad (6.47)$$

Since $K \gtrsim M^{1-\sigma}$ and $r - 2s \approx -1/4$, the gain in (6.47) clearly beats the 2δ -loss in (6.44).

Step 3: The error term. Using the commutator estimate (Lemma 2.10), we have that

$$\|\mathcal{E}_{K,i}^k\|_{C_u^\eta C_v^{r-1}} \lesssim K^{r-1-s+\eta+2\delta} \|\mathbb{I}_{ij}^{\diamond,k}(\phi)\|_{C_u^s C_v^s} \|\partial_v \zeta^j\|_{C_u^{1-r-2\delta} C_v^{r-1}}. \quad (6.48)$$

Similar as in Step 1 or 2, it is easy to see for ζ of type (+), (+-), or (S) that

$$\|\partial_v \zeta^j\|_{C_u^{1-r-2\delta} C_v^{r-1}} \lesssim \mathcal{G}(\zeta)\theta. \quad (6.49)$$

Together with the bilinear estimate (Proposition 2.5), we obtain from (6.48) and (6.49) that

$$\begin{aligned} \|\mathcal{E}_{K,i}^k \partial_u (A_{M,m}^{+,i} \phi_M^{+,m})\|_{C_u^{r-1} C_v^{r-1}} &\lesssim \|\mathcal{E}_{K,i}^k\|_{C_u^\eta C_v^{r-1}} \|\partial_u (A_{M,m}^{+,i} \phi_M^{+,m})\|_{C_u^\eta C_v^s} \\ &\lesssim K^{r-1-s+\eta+2\delta} M^{1-s+\eta} \mathcal{G}(\zeta)\theta^3. \end{aligned}$$

Since $K \gtrsim M^{1-\sigma}$, $r-1-s+\eta+2\delta \approx -3/4$, and $1-s+\eta \approx 1/2$, this term is acceptable. \square

We now extend the estimate of the (+-)-(S)-case from the null form to the full nonlinearity. As described in Sect. 4.3, the only problematic term is the unfortunate resonance.

Proposition 6.6 (Resonant term). *Let Hypothesis 3.8 be satisfied, let $1 \leq k \leq D$, and let M_1 and M_2 be frequency scales satisfying $M_1 \sim_\delta M_2$. Then, it holds that*

$$\begin{aligned} &\|\mathbb{I}_{ij}^{\diamond,k}(\phi) \partial_u (B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2}) \partial_v \psi^j \\ &- \left(\mathbb{I}_{ij}^{\diamond,k}(\phi) B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \|_{C_u^{r-1} C_v^{r-1}} \lesssim (M_1 M_2)^{-\eta} \theta^4. \end{aligned} \quad (6.50)$$

Furthermore, the resonant part satisfies

$$\|\mathbb{I}_{ij}^{\diamond,k}(\phi) B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j)\|_{C_u^s C_v^{r-1}} \lesssim M_2^{1-r-s+2\eta} \theta^3. \quad (6.51)$$

Proof. We first prove (6.50). To this end, we decompose

$$\mathbb{I}_{ij}^{\diamond,k}(\phi) \partial_u (B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2}) \partial_v \psi^j \quad (6.52)$$

$$\begin{aligned} &- \left(\mathbb{I}_{ij}^{\diamond,k}(\phi) B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \\ &= \left[\mathbb{I}_{ij}^{\diamond,k}(\phi) \left(\partial_u (B_{M_1, M_2, m_1 m_2}^i \phi_{M_1}^{+,m_1} \phi_{M_2}^{-,m_2}) \partial_v \psi^j \right. \right. \end{aligned} \quad (6.53)$$

$$\begin{aligned} &- \left. \left(B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \right] \\ &+ \left[\mathbb{I}_{ij}^{\diamond,k}(\phi) \left(\left(B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \right) \right. \\ &- \left. \left(\mathbb{I}_{ij}^{\diamond,k}(\phi) B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-,m_2} \downarrow_\sigma^v \partial_v \psi^j) \right) \otimes_\sigma^u \partial_u \phi_{M_1}^{+,m_1} \right]. \end{aligned} \quad (6.54)$$

We estimate the two terms (6.53) and (6.54) separately. The first term (6.53) can be controlled using the multiplication estimate (Corollary 2.8) and Proposition 5.4. Using Corollary 2.11 and Proposition 5.4, we have that

$$\begin{aligned} & \left\| \mathbb{I}_{ij}^{\diamond,k}(\phi) \left(B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-, m_2} \downarrow_{\sigma}^v \partial_v \psi^j) \right) \otimes_{\sigma}^u \partial_u \phi_{M_1}^{+, m_1} \right\| \\ & - \left(\mathbb{I}_{ij}^{\diamond,k}(\phi) B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-, m_2} \downarrow_{\sigma}^v \partial_v \psi^j) \right) \otimes_{\sigma}^u \partial_u \phi_{M_1}^{+, m_1} \Big\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \left\| \mathbb{I}_{ij}^{\diamond,k}(\phi) \right\|_{C_u^s C_v^s} \left\| B_{M_1, M_2, m_1 m_2}^i (\phi_{M_2}^{-, m_2} \downarrow_{\sigma}^v \partial_v \psi^j) \right\|_{C_u^s C_v^{r-1}} \left\| \partial_u \phi_{M_1}^{+, m_1} \right\|_{C_u^{s-1+2\delta}} \\ & \lesssim M_1^{2\delta} M_2^{1-s-r} \theta^4. \end{aligned}$$

Since $1 - s - r \approx -1/4$, this term is acceptable. The remaining estimate (6.51) for the resonant term follows directly from the multiplication estimate (Corollary 2.8) and Proposition 5.4. \square

6.3. The remainder terms. It remains to extend the \checkmark -estimates from the null form to the full nonlinearity.

Proposition 6.7. *Let Hypothesis 3.8 be satisfied, let $1 \leq k \leq D$, and let $\zeta_1, \zeta_2: \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ satisfy*

$$\text{Type}(\zeta_1, \zeta_2) \in \{(-)-(+), (-)-(+ -), (-)-(S), (+-)-(+ -), (S)-(S)\}.$$

Then, it holds that

$$\left\| \mathbb{I}_{ij}^{\diamond,k}(\phi) \partial_u \zeta_1^i \partial_v \zeta_2^j \right\|_{C_u^{r-1} C_v^{r-1}} \lesssim \mathcal{G}(\zeta_1) \mathcal{G}(\zeta_2) \theta^2.$$

Proof. Using the multiplication estimate (Corollary 2.8) and the composition estimate (Lemma 2.9), we obtain that

$$\begin{aligned} \left\| \mathbb{I}_{ij}^{\diamond,k}(\phi) \partial_u \zeta_1^i \partial_v \zeta_2^j \right\|_{C_u^{r-1} C_v^{r-1}} & \lesssim \left\| \mathbb{I}_{ij}^{\diamond,k}(\phi) \right\|_{C_u^s C_v^s} \left\| \partial_u \zeta_1^i \partial_v \zeta_2^j \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim (1 + \|\phi\|_{C_u^s C_v^s})^{10} \left\| \partial_u \zeta_1^i \partial_v \zeta_2^j \right\|_{C_u^{r-1} C_v^{r-1}} \\ & \lesssim \left\| \partial_u \zeta_1^i \partial_v \zeta_2^j \right\|_{C_u^{r-1} C_v^{r-1}}. \end{aligned}$$

The desired estimate now follows directly from Proposition 5.5. \square

6.4. Proof of Theorem 6.1. By combining the earlier propositions, we now prove the main theorem of this section.

Proof of Theorem 6.1. It only remains to combine our previous estimates. To this end, we decompose

$$\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u \phi^i \partial_v \phi^j = \mathbb{I}_{ij}^{\diamond,a}(\phi) \sum_{\substack{\text{Type } \zeta_1 = \\ (+), (-), (+-), (S)}} \sum_{\substack{\text{Type } \zeta_2 = \\ (-), (+), (+-), (S)}} \partial_u \zeta_1^i \partial_v \zeta_2^j.$$

Before continuing with the proof, we encourage the reader to review Fig. 2, which gives an overview of the relevant cases.

Case (+)-(-): In this case, $\zeta_1^i = A_{M,m}^{+,i} \phi_M^{+,m}$ and $\zeta_2^j = A_{N,n}^{-,j} \phi_N^{-,n}$ for some $M, N \geq 1$. If $N \leq M^{1-\delta}$, we use Proposition 6.2, which contributes (3.29) in F^+ . If $M \sim_\delta N$, we use Proposition 6.3, which contributes $-G_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}$, (3.31) in F^+ , and (3.35) in F^- . If $M \leq N^{1-\delta}$, which use the symmetry in the u and v -variables and Proposition 6.2, which contributes (3.33) in F^- .

The cases (+)-(+), (+)-(+), (+)-(S), (-)-(-), (+)-(-), and (S)-(-): Using the symmetry in the u and v -variables, all six cases are addressed by Proposition 6.5. The corresponding contributions to F^\pm are (3.30) and (3.34).

The cases (+-)-(S) and (S)-(+): Using the symmetry in the u and v -variables, both cases are addressed by Proposition 6.6. The corresponding contributions to F^\pm are (3.32) and (3.36).

The remaining cases: Using the symmetry in the u and v -variables, all remaining cases are addressed by Proposition 6.7. The corresponding contribution only enters into the remainder. \square

7. Duhamel Integral Approximation

In Sect. 6, we obtained a detailed description of the nonlinearity $\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u \phi^i \partial_v \phi^j$. In order to close the fixed point arguments for A^+ , A^- , and ψ , however, we need a description of the localized Duhamel integral

$$\chi^+ \chi^- \text{Duh}[\mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_u \phi^i \partial_v \phi^j]. \quad (7.1)$$

From Proposition 2.23, we already know that the Duhamel integral maps the remainder \mathcal{E} in Theorem 6.1 into $C_u^r C_v^r$, which is an acceptable error. It therefore remains to analyze the Duhamel integral of the three structured components in Theorem 6.1. The main result of this section is contained in the following proposition.

Proposition 7.1 (Duhamel integral approximation). *Let Hypothesis 3.8 be satisfied and let $1 \leq a \leq D$. Then, it holds that*

$$-\chi^+ \chi^- \text{Duh}[\mathbb{I}_{ij}^a(\phi) \partial_u \phi^i \partial_v \phi^j] = \sum_{M \sim_\delta N} G_{\chi,M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n} \quad (7.2)$$

$$- \sum_M P_{\leq M^{1-\sigma}}^u \left(\chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right) \phi_M^{+,m}(u) \quad (7.3)$$

$$+ \sum_N P_{\leq N^{1-\sigma}}^v \left(\chi^+ \chi^- \int_u^v du' P_{\leq N^{1-\sigma}}^v F_{N,n}^{-,a}(u', v) \right) \phi_N^{-,n}(v) \quad (7.4)$$

$$+ \mathcal{R}^a, \quad (7.5)$$

where the remainder satisfies

$$\|\mathcal{R}^a\|_{C_u^r C_v^r} \lesssim \theta^3.$$

The additional frequency projections outside the integrals in (7.3) and (7.4) may seem superfluous. However, they are necessary due to the u -dependence of $\int_u^v dv'(\dots)$, v -dependence of $\int_u^v du'(\dots)$, and the cut-off functions χ^+ and χ^- .

7.1. The low \times high-term. In this section, we analyze the contribution of $F_{M,m}^{+,a} \otimes_{\sigma}^u \partial_u \phi_M^{+,m}$ to the Duhamel integral. The desired estimate is included in the next lemma, which is the main estimate in this section.

Lemma 7.2 (Duhamel integral approximation for the low \times high-term). *Let Hypothesis 3.8 be satisfied, let $1 \leq a \leq D$, and let M be a frequency-scale. Then, it holds that*

$$\begin{aligned} & \left\| \chi^+ \chi^- \text{Duh} \left[F_{M,m}^{+,a} \otimes_{\sigma}^u \partial_u \phi_M^{+,m} \right] - P_{\leq M^{1-\sigma}}^u \left(\chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right) \phi_M^{+,m} \right\|_{C_u^r C_v^r} \\ & \lesssim M^{-\eta} \theta^3. \end{aligned} \quad (7.6)$$

Before we can prove Lemma 7.2, we prove two auxiliary lemmas.

Lemma 7.3 (Structured representation of Duhamel integral I). *Let $1 \leq m \leq D$, let $M \geq 1$, let $\zeta \in C_u^s C_v^{r-1}$. Then, it holds that*

$$\text{Duh} [\zeta \partial_u \phi_M^{+,m}] = \left(\int_u^v dv' \zeta(u, v') \right) \phi_M^{+,m} - I \left[\text{Tr} (\zeta \phi_M^{+,m}) \right] \Big|_u^v - \text{Duh} [\partial_u \zeta \phi_M^{+,m}]. \quad (7.7)$$

Proof. Using Proposition 2.23 and integration by parts, it holds that

$$\begin{aligned} & \text{Duh} [\zeta \partial_u \phi_M^{+,m}] \\ &= - \int_u^v dv' \int_u^{v'} du' \zeta(u', v') \partial_u \phi_M^{+,m}(u') \\ &= - \int_u^v dv' \left(\zeta(u', v') \phi_M^{+,m}(u') \Big|_{u'=u}^{v'} \right) + \int_u^v dv' \int_u^{v'} du' \partial_u \zeta(u', v') \phi_M^{+,m}(u') \\ &= \left(\int_u^v dv' \zeta(u, v') \right) \phi_M^{+,m}(u) - \int_u^v dv' \zeta(v', v') \phi_M^{+,m}(v') + \int_u^v dv' \int_u^{v'} du' \partial_u \zeta(u', v') \phi_M^{+,m}(u'). \end{aligned}$$

Using the definitions of the integral operator I and trace operator Tr , this yields (7.7). \square

In the proof of Proposition 7.1, the first summand in (7.7) will be responsible for the non-perturbative term (7.3). In order to estimate the second and third summand in (7.7), we prove the following estimate for $F_{M,m}^{+,a}$.

Lemma 7.4 (Bounds on $F_{M,m}^{+,a}$). *Let Hypothesis 3.8 be satisfied, let $1 \leq a, m \leq D$, and let M be a frequency scale. Then, it holds that*

$$\| P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \|_{C_u^s C_v^{r-1}} \lesssim M^{r-s+2\eta} \theta^2, \quad (7.8)$$

$$\| \text{Tr} (F_{M,m}^{+,a} \otimes_{\sigma}^u \phi_M^{+,m}) \|_{C_x^{r-1}} \lesssim M^{-\eta} \theta^3. \quad (7.9)$$

Proof. We first recall that $F_{M,m}^{+,a}$ consists of four different summands, which are given by (3.29), (3.30), (3.31), and (3.32). We now prove the two estimates (7.8) and (7.9) simultaneously, but distinguish between the four summands in $F_{M,m}^{+,a}$.

Case 1: Estimate for (3.29). We first control the contribution to (7.8). Using the bilinear estimate, we have that

$$\begin{aligned} & \| A_{M,m}^{+,i} \Pi_{ij}^a(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \|_{C_u^s C_v^r} \lesssim \| A_{M,m}^{+,i} \|_{C_u^s C_v^r} \| \Pi_{ij}^a(\phi) \|_{C_u^s C_v^r} \| \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \|_{C_u^s C_v^{r-1}} \\ & \lesssim N^{r-s} \theta^3. \end{aligned}$$

After summing over $N \lesssim M^{1-\delta}$, this yields an acceptable contribution to (7.8). We now turn to (7.9). Using the trace estimate (Lemma 2.21), we have that

$$\begin{aligned} & \left\| \operatorname{Tr} \left(\left(A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \phi_M^{+,m} \right) \right\|_{C_x^{r-1}} \\ & \lesssim \left\| \left(A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \right) \otimes_{\sigma}^u \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^{1-r'}} \\ & \lesssim \|A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi)\|_{C_u^s C_v^s} \|\partial_v (A_{N,n}^{-,j} \phi_N^{-,n})\|_{C_u^s C_v^{1-r'}} \|\phi_M^{+,m}\|_{C_u^{r-1}} \\ & \lesssim N^{2-r'-s} M^{r-1-s} \theta^4. \end{aligned}$$

Since $N \leq M^{1-\delta}$, we have that

$$N^{2-r'-s} M^{r-1-s} \lesssim M^{(1-\delta)(2-r'-s)+r-1-s} \lesssim M^{1-2s+\eta-\delta/6}.$$

Since $\delta \gg \eta \gg 1 - 2s$, this is acceptable.

Case 2: Estimate for (3.30). We first decompose

$$\begin{aligned} & P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) \partial_v \zeta^j) \\ & = P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j) + P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\gg M^{1-\sigma}}^u \partial_v \zeta^j) \\ & = P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j) + P_{\leq M^{1-\sigma}}^u ([P_{\leq M^{1-\sigma}}^u, A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi)] P_{\gg M^{1-\sigma}}^u \partial_v \zeta^j) \end{aligned} \quad (7.10)$$

We estimate the contributions of the first and second summand in (7.10) separately.

We start with the first summand in (7.10). In order to prove (7.8) and (7.9) simultaneously, we let $\gamma \in \{1 - r', s\}$. Then, it holds that

$$\begin{aligned} & \|P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j)\|_{C_u^\gamma C_v^{r-1}} \\ & \lesssim \|A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j\|_{C_u^\gamma C_v^{r-1}} \\ & \lesssim \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\Pi_{ij}^{\circ,a}(\phi)\|_{C_u^s C_v^s} \|P_{\leq M^{1-\sigma}}^u \partial_v \zeta\|_{C_u^\gamma C_v^{r-1}}. \end{aligned} \quad (7.11)$$

By inserting the types (+), (+−), and (S), it is easy to see that

$$\|P_{\leq M^{1-\sigma}}^u \partial_v \zeta\|_{C_u^{1-r'} C_v^{r-1}} \lesssim M^\delta \mathcal{G}(\zeta) \theta \quad \text{and} \quad \|P_{\leq M^{1-\sigma}}^u \partial_v \zeta\|_{C_u^s C_v^{r-1}} \lesssim M^{r-s+2\eta} \mathcal{G}(\zeta) \theta.$$

Inserting this into (7.11), we obtain the two inequalities

$$\|P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j)\|_{C_u^{1-r'} C_v^{r-1}} \lesssim M^\delta \mathcal{G}(\zeta) \theta^2, \quad (7.12)$$

$$\|P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j)\|_{C_u^s C_v^{r-1}} \lesssim M^{r-s+2\delta} \mathcal{G}(\zeta) \theta^2. \quad (7.13)$$

Due to (7.13), we see that the contribution to (7.9) is acceptable. Using the trace estimate (Lemma 2.21) and (7.12), it also holds that

$$\begin{aligned} & \left\| \operatorname{Tr} \left(P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j) \phi_M^{+,m} \right) (x) \right\|_{C_x^{r-1}} \\ & \lesssim \|P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j) \phi_M^{+,m}\|_{C_u^{1-r'} C_v^{r-1}} \\ & \lesssim \|P_{\leq M^{1-\sigma}}^u (A_{M,m}^{+,i} \Pi_{ij}^{\circ,a}(\phi) P_{\leq M^{1-\sigma}}^u \partial_v \zeta^j)\|_{C_u^{1-r'} C_v^{r-1}} \|\phi_M^{+,m}\|_{C_u^{1-r'}} \\ & \lesssim M^{\delta+1-r'-s} \theta^2. \end{aligned} \quad (7.14)$$

Since $\delta + 1 - r' - s \approx -1/4$, this is acceptable.

We now turn to the second summand in (7.10). Using the commutator estimate (Lemma 2.10), we have that

$$\begin{aligned}
 & \left\| P_{\leq M^{1-\sigma}}^u ([P_{\leq M^{1-\sigma}}^u, A_{M,m}^{+,i} \mathbb{I}_{ij}^{\diamond,a}(\phi)] P_{\gg M^{1-\sigma}}^u \partial_v \zeta^j) \right\|_{C_u^s C_v^{r-1}} \\
 & \lesssim M^{(1-\sigma)(s-\eta)} \left\| [P_{\leq M^{1-\sigma}}^u, A_{M,m}^{+,i} \mathbb{I}_{ij}^{\diamond,a}(\phi)] P_{\gg M^{1-\sigma}}^u \partial_v \zeta^j \right\|_{C_u^s C_v^{r-1}} \\
 & \lesssim M^{(1-\sigma)(s-\eta)} M^{(1-\sigma)(\eta-s)} M^{-2(1-\sigma)\eta} \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|\mathbb{I}_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|\partial_v \zeta^j\|_{C_u^{3\eta} C_v^s} \\
 & \lesssim M^{-\eta} \theta \|\partial_v \zeta^j\|_{C_u^{3\eta} C_v^s}.
 \end{aligned}$$

By inserting the types (+), (+−), and (S) into ζ , it is easy to see that

$$\|\partial_v \zeta^j\|_{C_u^{3\eta} C_v^s} \lesssim \mathcal{G}(\zeta) \theta.$$

Thus, this yields an acceptable contribution to (7.8). The contribution to (7.9) can be controlled using the trace estimate (Lemma 2.21) similar as in (7.14).

Case 3: Estimate for (3.31). To prove (7.8), we use Proposition 2.5.(iv), which yields

$$\begin{aligned}
 & \|(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^v \partial_v \phi_N^{-,n}\|_{C_u^s C_v^\eta} \\
 & \lesssim \|(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j})\|_{C_u^s C_v^s} \|\partial_v \phi_N^{-,n}\|_{C_u^s C_v^{\eta-s}} \\
 & \lesssim N^{1-2s+\eta} \|\mathbb{I}_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|A_{M,m}^{+,i}\|_{C_u^s C_v^s} \|A_{N,n}^{-,j}\|_{C_u^s C_v^s} \|\phi_N^{-,n}\|_{C_v^s} \\
 & \lesssim N^{1-2s+\eta} \theta^3.
 \end{aligned} \tag{7.15}$$

Since $N \lesssim M^{1-\delta}$ and $1 - 2s + \eta \ll r - s$, this yields an acceptable contribution to (7.8). For (7.9), we note that

$$\begin{aligned}
 & \left\| \text{Tr} \left(\left((\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^v \partial_v \phi_N^{-,n} \right) \otimes_\sigma^u \phi_M^{+,m} \right) (x) \right\|_{C_x^{r-1}} \\
 & \lesssim \left\| \left((\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^v \partial_v \phi_N^{-,n} \right) \otimes_\sigma^u \phi_M^{+,m} \right\|_{C_u^\eta C_v^\eta} \\
 & \lesssim M^{\eta-s} \|(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j}) \odot^v \partial_v \phi_N^{-,n}\|_{C_u^\eta C_v^\eta} \|\phi_M^{+,m}\|_{C_u^s}.
 \end{aligned}$$

By reusing (7.15), this yields an acceptable contribution to (7.9).

Case 4: Estimate for (3.32). Using the bilinear estimate (Proposition 2.5) and Lemma 4.5, we have that

$$\begin{aligned}
 & \|\mathbb{I}_{ij}^{\diamond,a}(\phi) B_{M,N,mn}^i (\phi_N^{-,n} \downarrow_\sigma^v \partial_v \psi^j)\|_{C_u^s C_v^s} \\
 & \lesssim \|\mathbb{I}_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|B_{M,N,mn}^i\|_{C_u^s C_v^s} \|\phi_N^{-,n} \downarrow_\sigma^v \partial_v \psi^j\|_{C_u^s C_v^s} \\
 & \lesssim N^{\sigma(s+1-r)} \|\mathbb{I}_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|B_{M,N,mn}^i\|_{C_u^s C_v^s} \|\phi_N^{-,n} \downarrow_\sigma^v \partial_v \psi^j\|_{C_u^s C_v^{r-1}} \\
 & \lesssim N^{\sigma(s+1-r)} N^{1-s-r} \theta^3.
 \end{aligned}$$

Since $\sigma(s+1-r) + 1 - s - r \approx -1/4$, this yields an acceptable contribution to (7.8). The contribution to (7.9) can then be estimated as in Case 3. \square

Equipped with Lemma 7.3 and Lemma 7.4, we are now ready to prove Lemma 7.2.

Proof of Lemma 7.2. Using Lemma 7.3, it holds that

$$\begin{aligned} & \chi^+ \chi^- \text{Duh} \left[F_{M,m}^{+,a} \otimes_{\sigma}^u \partial_u \phi_M^{+,m} \right] - P_{\leq M^{1-\sigma}}^u \left(\chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right) \phi_M^{+,m} \\ &= -\chi^+ \chi^- I \left[\text{Tr} \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \phi_M^{+,m} \right) \right] \Big|_{x=u}^v \end{aligned} \quad (7.16)$$

$$+ \chi^+ \chi^- \text{Duh} \left[\partial_u \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right) \phi_M^{+,m} \right] \quad (7.17)$$

$$+ P_{> M^{1-\sigma}}^u \left(\chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right) \phi_M^{+,m}. \quad (7.18)$$

We first estimate (7.16). To this end, we let $\tilde{\chi}$ be a fattened version of χ . Using Lemma 2.16 and Lemma 7.4, we have that

$$\begin{aligned} \left\| \chi^+ \chi^- I \left[\text{Tr} \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \phi_M^{+,m} \right) \right] \Big|_{x=u}^v \right\|_{C_u^r C_v^r} &\lesssim \left\| I \left[\tilde{\chi} \text{Tr} \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \phi_M^{+,m} \right) \right] \right\|_{C_x^r} \\ &\lesssim \left\| \text{Tr} \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \phi_M^{+,m} \right) \right\|_{C_x^{r-1}} \\ &\lesssim M^{-\eta} \theta^3. \end{aligned}$$

We now turn to (7.17). Using Proposition 2.23, Proposition 2.5.(iii), and Lemma 7.4, we have that

$$\begin{aligned} \left\| \chi^+ \chi^- \text{Duh} \left[\partial_u \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right) \phi_M^{+,m} \right] \right\|_{C_u^r C_v^r} &\lesssim \left\| \partial_u \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right) \phi_M^{+,m} \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim \left\| \partial_u \left(P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right) \right\|_{C_u^r C_v^{r-1}} \left\| \phi_M^{+,m} \right\|_{C_u^{r-1}} \\ &\lesssim M^{1+\eta-s} M^{r-1-s} \left\| P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right\|_{C_u^s C_v^{r-1}} \left\| \phi_M^{+,m} \right\|_{C_u^s} \\ &\lesssim M^{2r-3s+3\eta} \theta^3. \end{aligned}$$

Since $3s - 2r \gg \eta$, this is acceptable. It remains to estimate (7.18). Using Proposition 2.5.(ii), we have that

$$\begin{aligned} & \left\| P_{> M^{1-\sigma}}^u \left(\chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right) \phi_M^{+,m} \right\|_{C_u^r C_v^r} \\ &\lesssim \left\| \chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right\|_{C_u^r C_v^r} \left\| \phi_M^{+,m} \right\|_{C_u^{\sigma r}} \\ &\lesssim M^{\sigma r-s} \theta \left\| \chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right\|_{C_u^r C_v^r}. \end{aligned}$$

Using the trace and integral estimates (Lemma 2.21 and Lemma 2.19) as well as Lemma 7.4, we have that

$$\begin{aligned} M^{\sigma r-s} \theta \left\| \chi^+ \chi^- \int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u, v') \right\|_{C_u^r C_v^r} &\lesssim M^{\sigma r-s} \theta \left\| P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a} \right\|_{C_u^r C_v^{r-1}} \\ &\lesssim M^{\sigma r-s} M^{(1-\sigma)(r-s)} M^{r-s} \theta^3. \end{aligned}$$

Since $3s - 2r \gg \sigma$, this is acceptable. \square

7.2. The high \times high-term. We now analyze the contribution of $G_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}$ to the Duhamel estimate. The following lemma constitutes the main estimate of this section.

Lemma 7.5. (Duhamel integral approximation of the high \times high-term) *Let Hypothesis 3.8 be satisfied, let $1 \leq a \leq D$, and let M, N be frequency-scales satisfying $M \sim_\delta N$. Then, it holds that*

$$\|\chi^+ \chi^- \text{Duh} [G_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}] - G_{\chi,M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n}\|_{C_u^r C_v^r} \lesssim (MN)^{-\eta} \theta^3. \quad (7.19)$$

Here, $G_{\chi,M,N,mn}^a$ is as in (3.38). Before we proceed with the proof of Lemma 7.5, we prove the following algebraic lemma.

Lemma 7.6. *Let M, N be frequency-scales satisfying $M \sim_\delta N$. Then, it holds that*

$$\begin{aligned} & \chi^+ \chi^- \text{Duh} [G_{M,N,mn}^a \partial_u \phi_M^{+,m} \partial_v \phi_N^{-,n}] - G_{\chi,M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n} \\ &= (\chi^+ \chi^- G_{M,N,mn}^a - G_{\chi,M,N,mn}^a) \phi_M^{+,m} \phi_N^{-,n} \end{aligned} \quad (7.20)$$

$$- \chi^+ \chi^- \left[\text{Tr}_u \left(G_{M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n} \right) \right] \quad (7.21)$$

$$+ I_v \left[\partial_v G_{M,N,mn}^a \phi_N^{-,n} \right] (u, x) \Big|_{x=u}^v \phi_M^{+,m}(u) \quad (7.22)$$

$$+ I \left[\text{Tr} \left(G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right) \right] (x) \Big|_{x=u}^v \quad (7.23)$$

$$+ \text{Duh} \left[\partial_u G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right]. \quad (7.24)$$

Proof of Lemma 7.6. After applying Lemma 7.3 with $\zeta = G_{M,N,mn}^a \partial_v \phi_N^{-,n}$, it only remains to show that

$$\begin{aligned} & \left(\int_u^v dv' G_{M,N,mn}^a(u, v') \partial_v \phi_N^{-,n}(v') \right) \phi_M^{+,m}(u) \\ &= G_{M,N,mn}^a(u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v) - \text{Tr}_u \left(G_{M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n} \right) (u, v) \\ & - I_v \left[\partial_v G_{M,N,mn}^a \phi_N^{-,n} \right] (u, x) \Big|_{x=u}^v \phi_M^{+,m}(u). \end{aligned}$$

This follows directly from an integration by parts. \square

Proof of Lemma 7.5. Throughout the proof, we let $\tilde{\chi}$, $\tilde{\chi}^+$, and $\tilde{\chi}^-$ be fattened versions of χ , χ^+ , and χ^- , respectively. Using Lemma 7.6, it remains to estimate (7.20), (7.21), (7.22), (7.23), and (7.24) in $C_u^r C_v^r$.

Estimate of (7.20): From the definition of $G_{\chi,M,N,mn}^a$, it follows that

$$\begin{aligned} & (\chi^+ \chi^- G_{M,N,mn}^a - G_{\chi,M,N,mn}^a) \phi_M^{+,m} \phi_N^{-,n} \\ &= [\chi^+ \chi^-, P_{\leq M^{1-\sigma}}^u P_{\leq N^{1-\sigma}}^v] \left(\Pi_{ij}^a(\phi) A_{M,m}^{+,i} A_{N,n}^{+,j} \right) \phi_M^{+,m} \phi_N^{-,n}. \end{aligned}$$

Using $M \sim_\delta N$, this term can easily be controlled using Lemma 2.10.

Estimate of (7.21): We have that

$$\begin{aligned} \left\| \chi^+ \chi^- \operatorname{Tr}_u \left(G_{M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n} \right) \right\|_{C_u^r C_v^r} &\lesssim \max(M, N)^r \left\| G_{M,N,mn}^a \phi_M^{+,m} \phi_N^{-,n} \right\|_{L_{u,v}^\infty} \\ &\lesssim \max(M, N)^r M^{-s} N^{-s} \theta^3. \end{aligned}$$

Since $M \sim_\delta N$ and $r - 2s \approx -1/4$, this is acceptable.

Estimate of (7.22): Using the gain of a derivative through the integral, we have that

$$\begin{aligned} &\left\| \chi^+ \chi^- I_v \left[\partial_v G_{M,N,mn}^a \phi_N^{-,n} \right] (u, x) \right\|_{C_u^r C_v^r}^v \phi_M^{+,m}(u) \\ &\lesssim \max(M, N)^{2r} \left\| I_v \left[\tilde{\chi}^+ \tilde{\chi}^- \partial_v G_{M,N,mn}^a \phi_N^{-,n} \right] (u, x) \right\|_{C_u^r C_v^r}^v \phi_M^{+,m}(u) \\ &\lesssim \max(M, N)^{2r} \left\| I_v \left[\tilde{\chi}^+ \tilde{\chi}^- \partial_v G_{M,N,mn}^a \phi_N^{-,n} \right] (u, v) \right\|_{L_{u,v}^\infty} \left\| \phi_M^{+,m}(u) \right\|_{L_u^\infty} \\ &\lesssim \max(M, N)^{2r} N^{-1} N^{1-s} N^{-s} M^{-s} \theta^3 \\ &= \max(M, N)^{2r} M^{-s} N^{-2s} \theta^3. \end{aligned}$$

Since $3s - 2r \gg \delta$, this is acceptable.

Estimate of (7.23): Using Lemma 2.16, we have that

$$\begin{aligned} &\left\| \chi^+ \chi^- I \left[\operatorname{Tr} \left(G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right) \right] (x) \right\|_{C_u^r C_v^r}^v \\ &\lesssim \left\| I \left[\tilde{\chi} \operatorname{Tr} \left(G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right) \right] (x) \right\|_{C_x^r} \\ &\lesssim \left\| \operatorname{Tr} \left(G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right) \right\|_{C_x^{r-1}} \\ &\lesssim \left\| \operatorname{Tr} G_{M,N,mn}^a \right\|_{C_x^s} \left\| \phi_M^{+,m}(x) \partial_v \phi_N^{-,n}(x) \right\|_{C_x^{r-1}} \\ &\lesssim M^{-s} N^{r-s} \theta^3. \end{aligned}$$

Since $M \sim_\delta N$ and $r - 2s \approx -1/4$, this term is acceptable.

Estimate of (7.24): Using Proposition 2.23 and frequency-support considerations, it holds that

$$\begin{aligned} \left\| \chi^+ \chi^- \operatorname{Duh} \left[\partial_u G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right] \right\|_{C_u^r C_v^r} &\lesssim \left\| \partial_u G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right\|_{C_u^{r-1} C_v^{r-1}} \\ &\lesssim (MN)^{r-1} \left\| \partial_u G_{M,N,mn}^a \phi_M^{+,m} \partial_v \phi_N^{-,n} \right\|_{L_{u,v}^\infty} \\ &\lesssim (MN)^{r-1} M^{1-s} M^{-s} N^{1-s} \theta^3 \\ &\lesssim M^{r-2s} N^{r-s} \theta^3. \end{aligned}$$

Since $M \sim_\delta N$ and $3s - 2r \gg \delta$, this term is acceptable. \square

7.3. Proof of Proposition 7.1.

Proof of Proposition 7.1. Using Theorem 6.1, we have that

$$\begin{aligned} & -\chi^+\chi^-\text{Duh}\left[\mathbb{I}_{ij}^a(\phi)\partial_u\phi^i\partial_v\phi^j\right] \\ & = \chi^+\chi^-\sum_{M\sim_\delta N}\text{Duh}\left[G_{M,N,mn}^a\partial_u\phi_M^{+,m}\partial_v\phi_N^{-,n}\right] \end{aligned} \quad (7.25)$$

$$-\chi^+\chi^-\sum_M\text{Duh}\left[F_{M,m}^{+,a}\otimes_\sigma^u\partial_u\phi_M^{+,m}\right]-\chi^+\chi^-\sum_N\text{Duh}\left[F_{N,n}^{-,a}\otimes_\sigma^v\partial_v\phi_N^{-,n}\right] \quad (7.26)$$

$$+\chi^+\chi^-\text{Duh}\left[\mathcal{E}^a\right], \quad (7.27)$$

where the error term \mathcal{E} satisfies $\|\mathcal{E}\|_{C_u^{-1}C_v^{-1}} \lesssim \theta^2$. The contributions (7.25), (7.26), and (7.27) can all be addressed using previous lemmas.

Contribution of (7.25): This term is addressed in Lemma 7.5. It yields the first non-perturbative term (7.2) and contributes to the remainder \mathcal{R} .

Contribution of (7.26): Using the symmetry in the u and v -variables, both summands in (7.26) can be treated using Lemma 7.2. It yields the two non-perturbative terms (7.3) and (7.4) and contributes to the remainder \mathcal{R} .

Contribution of (7.27): Using Proposition 2.23, this term only contributes to the remainder \mathcal{R} . \square

8. Modulation Equations

In the previous section, we decomposed the Duhamel integral of the nonlinearity into the three structured components (7.2), (7.3), and (7.4) and the nonlinear remainder (7.5). As described in Sect. 9 below, the nonlinear remainder (7.5) will be absorbed into the ψ -portion of our Ansatz. The three structured components (7.2), (7.3), and (7.4), which cannot be treated perturbatively, are eliminated through the modulation equations. The modulation equations, which were previously stated in Definition 3.13, are given by

$$A_{M,m}^{+,a}(u,v) = \theta\delta_m^a - P_{\leq M^{1-\sigma}}^u \chi^+\chi^-\int_u^v dv' P_{\leq M^{1-\sigma}}^u F_{M,m}^{+,a}(u,v'), \quad (8.1)$$

$$A_{N,n}^{-,a}(u,v) = \theta\delta_n^a + P_{\leq N^{1-\sigma}}^v \chi^+\chi^-\int_u^v du' P_{\leq N^{1-\sigma}}^v F_{N,n}^{-,a}(u',v), \quad (8.2)$$

$$B_{M,N,mn}^b(u,v) = 1\{M\sim_\delta N\}G_{\chi,M,N,mn}^b. \quad (8.3)$$

Here, F^+ , F^- , and G are as in Definition 3.12. We also recall that $\theta\delta_m^a$ and $\theta\delta_n^a$ in (8.1) and (8.2) are a result of the linear evolution.

The goal of this section is to prove the local well-posedness of the modulation equation, i.e., Proposition 3.14. As described in the introduction (Sect. 1.3), the modulation equations cannot be treated using classical ODE methods. Instead, we rely on the para-controlled approach of Gubinelli, Imkeller, and Perkowski [GIP15]. To this end, we now derive para-controlled versions of the modulation equations.

8.1. Para-controlled modulation equations. We make the para-controlled Ansatz

$$A_{M,m}^{+,a} = \sum_{N \leq M^{1-\delta}} X_{M,N,mn}^{+,a} \otimes^v \phi_N^{-,n} + Y_{M,m}^{+,a}, \quad (8.4)$$

$$A_{N,n}^{-,a} = \sum_{M \leq N^{1-\delta}} X_{N,M,nm}^{-,a} \otimes^u \phi_M^{+,m} + Y_{N,n}^{-,a}. \quad (8.5)$$

To guide the reader, we mention that our analysis below yields uniform bounds for

$$X_{M,N}^+, X_{N,M}^- \in C_u^s C_v^s, \quad Y_M^+ \in C_u^s C_v^r, \quad \text{and} \quad Y_N^- \in C_u^r C_v^s.$$

Motivated by the frequency-support conditions for A^+ , A^- , and B (Conditions 3.5), we impose the following frequency-support conditions on the para-controlled modulations.

Condition 8.1 (Frequency-support conditions for X^\pm and Y^\pm). *For all $1 \leq i, m, n \leq D$ and all frequency scales M and N , we impose that*

$$P_{\gg M^{1-\sigma}}^u X_{M,N,mn}^{+,i} = 0, \quad (8.6)$$

$$P_{\gg M^{1-\sigma}}^u Y_{M,m}^{+,i} = 0, \quad (8.7)$$

$$P_{\gg N^{1-\sigma}}^v X_{N,M,nm}^{-,i} = 0, \quad (8.8)$$

$$P_{\gg N^{1-\sigma}}^v Y_{N,n}^{-,i} = 0. \quad (8.9)$$

We now state the para-controlled modulation equations for $X_{M,N}^+$, $X_{N,M}^-$, Y_M^+ , and Y_N^- , which will be derived momentarily (see Proposition 8.3).

Definition 8.2 (*Para-controlled modulation equations*). For all $1 \leq a, m, n \leq D$ and frequency scales M and N , the para-controlled modulation equations are given by

$$X_{M,N,mn}^{+,a} = -P_{\leq M^{1-\sigma}}^u \left(\chi^+ \chi^- \mathcal{X}_{M,N,mn}^{+,a} \right) \quad (8.10)$$

and

$$Y_{M,m}^{+,a} = \theta \delta_m^a - P_{\leq M^{1-\sigma}}^u \text{Com}_{\chi, \otimes, I}^v (\mathcal{X}_{M,N,mn}^{+,a}, \phi_N^{-,n}) \quad (8.11)$$

$$- P_{\leq M^{1-\sigma}}^u \left(\chi^+ \chi^- \left(I_v \mathcal{Y}_{M,m}^{+,a}(u, v) - I_v \mathcal{Y}_{M,m}^{+,a}(u, u) \right) \right). \quad (8.12)$$

Here, $\text{Com}_{\chi, \otimes, I}^v$ is as in Definition 2.24. Furthermore, $\mathcal{X}_{M,N,mn}^{+,a}$ is given by

$$\mathcal{X}_{M,N,mn}^{+,a} = P_{\leq M^{1-\sigma}}^u \left[\mathbb{I}_{ij}^a(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right. \quad (8.13)$$

$$+ \sum_{\substack{L: \\ N^{\frac{1}{1-\delta}} \leq L \ll M}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} \phi_L^{+,i} X_{L,N,\ell n}^{+,j} \quad (8.14)$$

$$\left. + \sum_{\substack{K: \\ K \sim_\delta N \\ K \leq M^{1-\delta}}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} B_{K,N,kn}^j \phi_K^{+,k} \right]. \quad (8.15)$$

Finally, the driving force $\mathcal{Y}_{M,m}^{+,a}$ is given by

$$\mathcal{Y}_{M,m}^{+,a} = \mathcal{Y}_{M,m}^{(1),+,a} + \mathcal{Y}_{M,m}^{(2),+,a} + \mathcal{Y}_{M,m}^{(3),+,a} + \mathcal{Y}_{M,m}^{(4),+,a}, \quad (8.16)$$

where the four summands are given by

$$\mathcal{Y}_{M,m}^{(1),+,a} = P_{\leq M^{1-\sigma}}^u \left[\sum_{N \leq M^{1-\delta}} \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right) \otimes^v \partial_v \phi_N^{-,n} \right. \quad (8.17)$$

$$+ \sum_{L \ll M} \sum_{N \leq L^{1-\delta}} X_{L,N,\ell n}^{+,j} \phi_L^{+, \ell} \left(\left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \right) \odot^v \partial_v \phi_N^{-,n} \right) \quad (8.18)$$

$$+ \sum_{\substack{K,N: \\ K,N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \phi_K^{+,k} \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} B_{K,N,kn}^j \right) \otimes^v \partial_v \phi_N^{-,n} \quad (8.19)$$

$$+ \sum_{\substack{N: \\ N \sim_{\delta} M}} \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right) \odot^v \partial_v \phi_N^{-,n} \Big], \quad (8.20)$$

$$\mathcal{Y}_{M,m}^{(2),+,a} = P_{\leq M^{1-\sigma}}^u \left[\sum_{L \gtrsim M} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \partial_v A_{L,\ell}^{+,j} \right. \quad (8.21)$$

$$+ \sum_{\substack{K,N: \\ \max(K,N) > M^{1-\delta}, \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v \left(B_{K,N,kn}^j \phi_K^{+,k} \phi_N^{-,n} \right) \Big], \quad (8.22)$$

$$\mathcal{Y}_{M,m}^{(3),+,a} = P_{\leq M^{1-\sigma}}^u \left[\sum_{L \ll M} \sum_{N \leq L^{1-\delta}} \phi_L^{+, \ell} \text{Com}_{\otimes}^v \left(\mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} X_{L,N,\ell n}^{+,j}, \partial_v \phi_N^{-,n} \right) \right], \quad (8.23)$$

$$\mathcal{Y}_{M,m}^{(4),+,a} = P_{\leq M^{1-\sigma}}^u \left[\sum_{N \leq M^{1-\delta}} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v A_{N,n}^{-,j} \phi_N^{-,n} \right. \quad (8.24)$$

$$+ \sum_{L \ll M} \sum_{N \leq L^{1-\delta}} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \left(\partial_v X_{L,N,\ell n}^{+,j} \otimes^v \phi_N^{-,n} \right) \quad (8.25)$$

$$+ \sum_{L \ll M} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \partial_v Y_{L,\ell}^{+,j} \quad (8.26)$$

$$+ \sum_{\substack{K,N: \\ K,N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v \left(B_{K,N,kn}^j \right) \phi_K^{+,k} \phi_N^{-,n} \quad (8.27)$$

$$+ A_{M,m}^{+,i} \mathbb{I}_{ij}^{\diamond,a}(\phi) \partial_v \psi^j \quad (8.28)$$

$$+ \sum_{\substack{N: \\ N \sim_{\delta} M}} \mathbb{I}_{ij}^{\diamond,a}(\phi) B_{M,N,mn}^i \left(\phi_N^{-,n} \odot_{\sigma}^v \partial_v \psi^j \right) \Big]. \quad (8.29)$$

The para-controlled modulation equations for $X_{N,M,nm}^{-,a}$ and $Y_{N,n}^{-,a}$ are similar to (8.10)–(8.29) but with reversed roles of the u and v -variables.

In the definition of $\mathcal{Y}_{M,m}^{+,a}$, we have collected similar terms in four different groups. The individual components $\mathcal{Y}^{(j)}$ contain the following terms:

- (1) The terms in $\mathcal{Y}^{(1)}$ all contain $\phi_N^{-,n}$ but involve a second factor at a comparable or higher v -frequency.
- (2) The terms in $\mathcal{Y}^{(2)}$ contains at least two factors at u -frequencies $\gtrsim M^{1-\sigma}$.
- (3) The term in $\mathcal{Y}^{(3)}$ has a commutator structure.
- (4) The terms in $\mathcal{Y}^{(4)}$ are “easy”.

Of course, the motivation behind (8.13)–(8.29) is not clear from the Definition, but can rather be seen from (the proof of) the following Proposition.

Proposition 8.3 (Para-controlled modulation equations). *If X^+ , X^- , Y^+ , and Y^- solve the system of para-controlled modulation equations (as in Definition 8.2), then A^+ and A^- solve the modulation equations (8.1) and (8.2).*

Proof. We only prove the result for A^+ , since the argument for A^- is similar. To this end, we fix a frequency-scale M . We now assign different parts of the right-hand side in (8.1) to either $X_{M,N}^+$, where $N \leq M^{1-\delta}$, or Y_M^+ . First, we recall from Definition 3.12 that F_M^+ consists of the four summands (3.29), (3.30), (3.31), and (3.32), which are treated separately.

Contribution of (3.29): We decompose

$$\begin{aligned} P_{\leq M^{1-\sigma}}^u \sum_{N \leq M^{1-\delta}} A_{M,m}^{+,i} \Pi_{ij}^{\diamond,a}(\phi) \partial_v (A_{N,n}^{-,j} \phi_N^{-,n}) \\ = P_{\leq M^{1-\sigma}}^u \sum_{N \leq M^{1-\delta}} \left(\Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right) \otimes^v \partial_v \phi_N^{-,n} \end{aligned} \quad (8.30)$$

$$+ P_{\leq M^{1-\sigma}}^u \sum_{N \leq M^{1-\delta}} \left(\Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} A_{N,n}^{-,j} \right) \otimes^v \partial_v \phi_N^{-,n} \quad (8.31)$$

$$+ P_{\leq M^{1-\sigma}}^u \sum_{N \leq M^{1-\delta}} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v A_{N,n}^{-,j} \phi_N^{-,n}. \quad (8.32)$$

The contribution of (8.30) is included in the commutator in (8.11) and (8.13). The contributions of (8.31) and (8.32) are contained in (8.17) and (8.24), respectively.

Contribution of (3.30): We distinguish between different types of ζ .

Contribution of (3.30), Type $\zeta = (+)$: We first decompose

$$\sum_L \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v (A_{L,\ell}^{+,j} \phi_L^{+, \ell}) = \sum_{L \ll M} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v A_{L,\ell}^{+,j} \phi_L^{+, \ell} \quad (8.33)$$

$$+ \sum_{L \gtrsim M} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \partial_v A_{L,\ell}^{+,j} \phi_L^{+, \ell}. \quad (8.34)$$

The contribution of (8.34) is included in (8.21). Thus, we now continue decomposing (8.33). Inserting the para-controlled Ansatz for $A_{L,\ell}^{+,j}$ from (8.4), we have that

$$\begin{aligned} \sum_{L: L \ll M} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \partial_v A_{L,\ell}^{+,j} \\ = \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \partial_v (X_{L,N,\ell n}^{+,j} \otimes^v \phi_N^{-,n}) + \sum_{L: L \ll M} \Pi_{ij}^{\diamond,a}(\phi) A_{M,m}^{+,i} \phi_L^{+, \ell} \partial_v Y_{L,\ell}^{+,j}. \end{aligned} \quad (8.35)$$

The second summand in (8.35) corresponds to (8.26) and it remains to treat the first summand in (8.35). Using the product rule, it holds that

$$\begin{aligned} & \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \phi_L^{+, \ell} \partial_v (X_{L, N, \ell n}^{+, j} \otimes^v \phi_N^{-, n}) \\ &= \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \phi_L^{+, \ell} (X_{L, N, \ell n}^{+, j} \otimes^v \partial_v \phi_N^{-, n}) \end{aligned} \quad (8.36)$$

$$+ \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \phi_L^{+, \ell} (\partial_v X_{L, N, \ell n}^{+, j} \otimes^v \phi_N^{-, n}). \quad (8.37)$$

The second summand (8.37) corresponds to (8.25). Therefore, it remains to treat the first summand (8.36). Using Definition 2.12, it holds that

$$\begin{aligned} & \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \phi_L^{+, \ell} (X_{L, N, \ell n}^{+, j} \otimes^v \partial_v \phi_N^{-, n}) \\ &= \sum_{N: N \leq M^{1-\delta}} \sum_{L: N^{1/(1-\delta)} \leq L \ll M} (\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \phi_L^{+, \ell} X_{L, N, \ell n}^{+, j}) \otimes^v \partial_v \phi_N^{-, n} \end{aligned} \quad (8.38)$$

$$+ \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} X_{L, N, \ell n}^{+, j} \phi_L^{+, \ell} \left((\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i}) \odot^v \partial_v \phi_N^{-, n} \right) \quad (8.39)$$

$$+ \sum_{\substack{L, N: \\ N^{1/(1-\delta)} \leq L \ll M}} \phi_L^{+, \ell} \text{Com}_{\otimes}^v \left((\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i}, X_{L, N, \ell n}^{+, j}, \partial_v \phi_N^{-, n}) \right). \quad (8.40)$$

The first summand (8.38) is contributes to the commutator in (8.11) and (8.14). The remaining summands (8.39) and (8.40) are responsible for (8.18) and (8.23), respectively.

Contribution of (3.30), Type $\zeta = (+-)$: The total contribution (without the $P_{\leq M^{1-\sigma}}^u$ operator) is given by

$$\begin{aligned} & \sum_{\substack{K, N: \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \partial_v (B_{K, N, kn}^j \phi_K^{+, k} \phi_N^{-, n}) \\ &= \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \partial_v (B_{K, N, kn}^j \phi_K^{+, k} \phi_N^{-, n}) \end{aligned} \quad (8.41)$$

$$+ \sum_{\substack{K, N: \\ \max(K, N) > M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \partial_v (B_{K, N, kn}^j \phi_K^{+, k} \phi_N^{-, n}). \quad (8.42)$$

The second summand (8.42) is responsible for (8.22). Using the product rule, we decompose (8.41) further as

$$\begin{aligned}
 & \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \partial_v (B_{K, N, kn}^j \phi_K^{+, k} \phi_N^{-, n}) \\
 &= \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} B_{K, N, kn}^j \phi_K^{+, k} \partial_v \phi_N^{-, n} \\
 &+ \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} \partial_v B_{K, N, kn}^j \phi_K^{+, k} \phi_N^{-, n}.
 \end{aligned} \tag{8.43}$$

The second summand in (8.43) is responsible for (8.27). The first summand in (8.43) is decomposed further by writing

$$\begin{aligned}
 & \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} B_{K, N, kn}^j \phi_K^{+, k} \partial_v \phi_N^{-, n} \\
 &= \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \left(\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} B_{K, N, kn}^j \phi_K^{+, k} \right) \ll^v \partial_v \phi_N^{-, n} \\
 &+ \sum_{\substack{K, N: \\ K, N \leq M^{1-\delta} \\ K \sim_{\delta} N}} \left(\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M, m}^{+, i} B_{K, N, kn}^j \phi_K^{+, k} \right) \gtrsim^v \partial_v \phi_N^{-, n}.
 \end{aligned}$$

The \ll^v -term contributes to the commutator in (8.11) and (8.15). The \gtrsim^v -term is included in included in (8.19).

Contribution of (3.30): Type $\zeta = (S)$: In this case, $\zeta = \psi$. The contribution is given by (4.18).

Cotribution of (3.31): This term coincides with (8.20).

Contribution of (3.32): This term coincides with (8.29). \square

8.2. Estimates for the modulation equations. In this section, we will derive the main estimates needed for the modulation equations. Similar as in Sects. 5–7, we capture the bounds used in our analysis using a single hypothesis (Hypothesis 8.4). This hypothesis will then form the basis of the contraction mapping argument in the proof of Proposition 3.14.

Hypothesis 8.4 (Modulation hypothesis). *We assume that*

$$\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta.$$

Furthermore, we assume that the modulations satisfy

$$\max \left(\sup_M \|A_M^+\|_{Mod_M^+}, \sup_N \|A_N^-\|_{Mod_N^-}, \sup_{M \sim_\delta N} \|B_{M,N}\|_{C_u^s C_v^s} \right) \leq 5\theta.$$

Finally, we assume that the para-controlled modulations satisfy

$$\max \left(\sup_{\substack{M,N: \\ N \leq M^{1-\delta}}} \|X_{M,N}^+\|_{C_u^s C_v^s}, \sup_{\substack{M,N: \\ M \leq N^{1-\delta}}} \|X_{N,M}^-\|_{C_u^s C_v^s}, \sup_M \|Y_M^+\|_{C_u^s C_v^r}, \sup_N \|Y_N^-\|_{C_u^r C_v^s} \right) \leq 2\theta.$$

Lemma 8.5 (Resonant-estimate). *Let Hypothesis 8.4 be satisfied, let $1 \leq i, m, n \leq D$, and let M and N be frequency scales. Then, it holds that*

$$\|A_{M,m}^{+,i} \odot^v \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta^2, \quad (8.44)$$

$$\|\mathbb{I}_{ij}^{\circ,a}(\phi) \odot^v \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta. \quad (8.45)$$

Proof. We prove (8.44) and (8.45) separately. The proof of (8.44) will be based on the para-controlled Ansatz (8.4), which was not used in the PDE-analysis. Once the first estimate (8.44) has been established, we prove the second estimate (8.45) by using (8.44) and our previous product estimate (Lemma 5.6).

Proof of (8.44): Recalling the para-controlled Ansatz (8.4), we have that

$$A_{M,m}^{+,i} = \sum_{K \leq M^{1-\delta}} X_{M,K,mk}^{+,i} \otimes^v \phi_K^{-,k} + Y_{M,m}^{+,i}.$$

We first estimate the contribution of the para-controlled terms $X_{M,K,mk}^{+,i} \otimes^v \phi_K^{-,k}$. To this end, we write

$$\begin{aligned} & (X_{M,K,mk}^{+,i} \otimes^v \phi_K^{-,k}) \odot^v \partial_v \phi_N^{-,n} \\ &= 1\{K \sim N\} \left(X_{M,K,mk}^{+,i} (\phi_K^{-,k} \odot^v \partial_v \phi_N^{-,n}) + \text{Com}_{\otimes, \odot}^v (X_{M,K,mk}^{+,i}, \phi_K^{-,k}, \partial_v \phi_N^{-,n}) \right), \end{aligned} \quad (8.46)$$

where the commutator is as in Definition 2.12. Using the multiplication estimate (Corollary 2.8) and Lemma 3.4, the first summand in (8.46) is bounded by

$$\begin{aligned} \|X_{M,K,mk}^{+,i} (\phi_K^{-,k} \odot^v \partial_v \phi_N^{-,n})\|_{C_u^s C_v^{r-1}} &\lesssim \|X_{M,K,mk}^{+,i}\|_{C_u^s C_v^s} \|\phi_K^{-,k} \odot^v \partial_v \phi_N^{-,n}\|_{C_v^{r-1}} \\ &\lesssim N^{r-2s} \theta^3. \end{aligned}$$

Since $r - 2s \approx -1/4$, this contribution is acceptable. Using Lemma 2.13, we have that

$$\begin{aligned} \|\text{Com}_{\otimes, \odot}^v (X_{M,K,mk}^{+,i}, \phi_K^{-,k}, \partial_v \phi_N^{-,n})\|_{C_u^s C_v^{r-1}} &\lesssim \|X_{M,K,mk}^{+,i}\|_{C_u^s C_v^s} \|\phi_K^{-,k}\|_{C_u^s C_v^s} \|\partial_v \phi_N^{-,n}\|_{C_u^s C_v^{-1+2\eta}} \\ &\lesssim N^{-s+2\eta} \theta^3. \end{aligned}$$

Since $-s + 2\eta \approx -1/2$, this term is acceptable. We now estimate the contribution of $Y_{M,m}^{+,i}$. Using the bilinear estimate (Proposition 2.5.(iv)), we have that

$$\begin{aligned} \|Y_{M,m}^{+,i} \odot^v \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} &\lesssim \|Y_{M,m}^{+,i}\|_{C_u^s C_v^r} \|\partial_v \phi_N^{-,n}\|_{C_v^{-r'}} \\ &\lesssim N^{1-s-r'} \theta^2. \end{aligned}$$

Since $1 - s - r' \approx -1/4$, this is acceptable.

Proof of (8.45): Using Lemma 2.14 and the composition estimate (Lemma 2.9), we have that

$$\begin{aligned} \|\mathbb{I}_{ij}^{\diamond,a}(\phi) \odot^v \partial_v \phi_N^{-,n} - \tilde{P}_N^v(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \\ \lesssim N^{r-1-s-(s-1)} \|\mathbb{I}_{ij}^{\diamond,a}(\phi)\|_{C_u^s C_v^s} \|\partial_v \phi_N^{-,n}\|_{C_v^{s-1}} \\ \lesssim N^{r-2s} \theta. \end{aligned}$$

Since $r - 2s \approx -1/4$, this is acceptable. As a result, it remains to prove

$$\|P_K^v(\mathbb{I}_{ij}^{\diamond,a}(\phi)) \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta^2$$

for all $K \sim N$. Using Bony's para-linearization (Lemma 2.9) and the bilinear estimate (Proposition 2.5), it further suffices to prove that

$$\|P_K^v \phi^k \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta^2$$

for all $K \sim N$. By inserting the Ansatz for ϕ , it then remains to prove the two estimates

$$\left\| \sum_M P_K^v (A_{M,m}^{+,k} \phi_M^{+,m}) \partial_v \phi_N^{-,n} \right\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta^3, \quad (8.47)$$

$$\sum_{\substack{\text{Type } \zeta = \\ (-), (+-), (S)}} \|P_K^v \zeta \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}} \lesssim N^{-\eta} \theta^2. \quad (8.48)$$

We start with the first estimate (8.47). Using Corollary 2.7, we have that

$$\left\| \sum_M P_K^v (A_{M,m}^{+,k} \phi_M^{+,m}) \partial_v \phi_N^{-,n} \right\|_{C_u^s C_v^{r-1}} \lesssim \sup_M \|\phi_M^{+,m}\|_{C_u^s C_v^s} \sup_M \|P_K^v A_{M,m}^{+,k} \partial_v \phi_N^{-,n}\|_{C_u^s C_v^{r-1}}.$$

The desired estimate then follows from (8.44). The second estimate (8.48) follows from the symmetry in the u and v -variables and Proposition 5.6, where $H_{M,m}^{+,i} = \delta_m^i$. \square

Lemma 8.6 (Estimate of $\mathcal{Y}^{(2)}$). *Let Hypothesis 8.4 be satisfied, let $1 \leq a, m \leq D$, and let M be a frequency scale. Then, it holds that*

$$\|\mathcal{Y}_{M,m}^{(2),a,+}\|_{C_u^s C_v^{r-1}} \lesssim M^{-\eta} \theta^3.$$

Proof. We estimate the two terms (8.21) and (8.22) separately. For (8.21), we have that

$$\begin{aligned}
& \left\| P_{\leq M^{1-\sigma}}^u \left(\sum_{L: L \gtrsim M} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} \partial_v A_{L,\ell}^{+,j} \phi_L^{+, \ell} \right) \right\|_{C_u^s C_v^{r-1}} \\
& \lesssim M^{(1-\sigma)s} \sum_{L: L \gtrsim M} \left\| P_{\gtrsim L}^u \left(\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} \partial_v A_{L,\ell}^{+,j} \right) \phi_L^{+, \ell} \right\|_{L_u^\infty C_v^{r-1}} \\
& \lesssim M^{(1-\sigma)s} \sum_{L: L \gtrsim M} L^{-2s} \left\| \mathbb{I}_{ij}^{\diamond, a}(\phi) \right\|_{C_u^s C_v^s} \left\| A_{M,m}^{+,i} \right\|_{C_u^s C_v^s} \left\| \partial_v A_{L,\ell}^{+,j} \right\|_{C_u^s C_v^{r-1}} \left\| \phi_L^{+, \ell} \right\|_{C_u^s} \\
& \lesssim M^{(1-\sigma)s} \sum_{L: L \gtrsim M} L^{r-3s} \theta^3 \\
& \lesssim M^{r-2s} \theta^3.
\end{aligned}$$

Since $r - 2s \approx -1/4$, this term is acceptable. We now estimate the second term in $\mathscr{Y}^{(2)}$, which is given by (8.22). To this end, we first note that $\max(K, N) > M^{1-\delta}$ and $K \sim_\delta N$ imply that $K, N \geq M^{1-2\delta}$. Then, we estimate

$$\begin{aligned}
& \left\| P_{\leq M^{1-\sigma}}^u \left[\sum_{\substack{K, N: \\ \max(K, N) > M^{1-\delta}, \\ K \sim_\delta N}} \mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} \partial_v \left(B_{K,N,kn}^j \phi_K^{+,k} \phi_N^{-,n} \right) \right] \right\|_{C_u^s C_v^{r-1}} \\
& \lesssim M^{(1-\sigma)s} \sum_{\substack{K, N: \\ K, N \geq M^{1-2\delta}, \\ K \sim_\delta N}} \left\| P_{\gtrsim K}^u \left(\mathbb{I}_{ij}^{\diamond, a}(\phi) A_{M,m}^{+,i} \partial_v \left(B_{K,N,kn}^j \phi_K^{+,k} \phi_N^{-,n} \right) \right) \right\|_{L_u^\infty C_v^{r-1}} \left\| \phi_K^{+,k} \right\|_{L_u^\infty} \\
& \lesssim M^{(1-\sigma)s} \sum_{\substack{K, N: \\ K, N \geq M^{1-2\delta}, \\ K \sim_\delta N}} K^{-2s} N^{r-s} \theta^4.
\end{aligned}$$

The dyadic sum can be estimated by

$$\begin{aligned}
M^{(1-\sigma)s} \sum_{\substack{K, N: \\ K, N \geq M^{1-2\delta}, \\ K \sim_\delta N}} K^{-2s} N^{r-s} & \lesssim M^{(1-\sigma)s} \sum_{\substack{N: \\ N \geq M^{1-2\delta}}} N^{r-s-(1-\delta)2s} \\
& \lesssim M^{(1-\sigma)s-(1-2\delta)(r-s-(1-\delta)2s)}.
\end{aligned}$$

Since $r - 2s \approx -1/4$, this is acceptable. \square

8.3. Proof of Proposition 3.14. Equipped with the estimates from Sect. 8.2, we can now prove the local well-posedness of the modulation equations.

Proof of Proposition 3.14. We use a contraction mapping argument. For expository purposes, we split the proof into three steps. In the first step, we set up the solution space \mathscr{S} , which captures the norms in Hypothesis 8.4, and the mapping Γ , which encodes the modulation equations. In the second step, we show that Γ maps a θ -ball in \mathscr{S} to itself. In

the third step, we briefly discuss the contraction property and the continuous dependence on the data $(\phi_0(0), \phi^+, \phi^-, \psi)$, but omit the details.

Step 1: Setup. We first define the norm

$$\begin{aligned} & \| (B, X^+, X^-, Y^+, Y^-) \|_{\mathcal{S}} \\ &= \max \left(5 \sup_{M \sim_{\delta} N} \|B_{M,N}\|_{C_u^s C_v^s}, 2 \sup_{\substack{M,N: \\ N \leq M^{1-\delta}}} \|X_{M,N}^+\|_{C_u^s C_v^s}, 2 \sup_{\substack{M,N: \\ M \leq N^{1-\delta}}} \|X_{N,M}^-\|_{C_u^s C_v^s}, \right. \\ & \quad \left. 2 \sup_M \|Y_M^+\|_{C_u^s C_v^r}, 2 \sup_N \|Y_N^-\|_{C_u^r C_v^s} \right). \end{aligned}$$

The absolute constants in $\|\cdot\|_{\mathcal{S}}$ have been chosen to match Hypothesis 8.4. We define the corresponding solution space \mathcal{S} as the set of functions that have finite \mathcal{S} -norm and satisfy the frequency-support conditions. More precisely, we define

$$\mathcal{S} := \left\{ (B, X^+, X^-, Y^+, Y^-) : \|(B, X^+, X^-, Y^+, Y^-)\|_{\mathcal{S}} < \infty \text{ and } \right. \\ \left. (3.15), (8.6), (8.7), (8.8), \text{ and } (8.9) \text{ hold} \right\}.$$

Furthermore, we define the θ -ball \mathcal{S}_{θ} by

$$\mathcal{S}_{\theta} := \left\{ (B, X^+, X^-, Y^+, Y^-) \in \mathcal{S} : \|(B, X^+, X^-, Y^+, Y^-)\|_{\mathcal{S}} \leq \theta \right\}.$$

For any element $(B, X^+, X^-, Y^+, Y^-) \in \mathcal{S}_{\theta}$, we define the corresponding modulations $A^+ = A^+[X^+, Y^+]$ and $A^- = A^-[X^-, Y^-]$ through the para-controlled Ansatz in (8.4) and (8.5). Using the bilinear estimate for families (Corollary 2.7), one easily obtains that

$$\begin{aligned} \sup_M \|A_M^+\|_{Mod_M^+} &\leq C\theta \sup_{M,N} \|X_{M,N}^+\|_{C_u^s C_v^s} + 2 \sup_M \|Y_M^+\|_{C_u^s C_v^r}, \\ \sup_N \|A_N^-\|_{Mod_N^-} &\leq C\theta \sup_{M,N} \|X_{N,M}^-\|_{C_u^s C_v^s} + 2 \sup_N \|Y_N^-\|_{C_u^r C_v^s}. \end{aligned}$$

In particular, the condition $(B, X^+, X^-, Y^+, Y^-) \in \mathcal{S}_{\theta}$ implies that A^+ and A^- satisfy the conditions in Hypothesis 8.4. We now define a map $\Gamma = \Gamma[B, X^+, X^-, Y^+, Y^-]$ which encodes the para-controlled modulation equations. We define

$$\begin{aligned} (\Gamma_B)_{M,N,mn}^b &:= \text{RHS of (3.41)}, \\ (\Gamma_X)_{M,N,mn}^{+,a} &:= \text{RHS of (8.10)}, \\ (\Gamma_Y)_{M,m}^{+,a} &:= \text{RHS of (8.11) + (8.12)}. \end{aligned}$$

The components $(\Gamma_X)^-$ and $(\Gamma_Y)^-$ are similar as $(\Gamma_X)^+$ and $(\Gamma_Y)^+$ but with reversed roles of the u and v -variables.

Step 2: Self-mapping property.

In this step, we prove that for any $(B, X^+, X^-, Y^+, Y^-) \in \mathcal{S}_{\theta}$, it holds that

$$\Gamma[B, X^+, X^-, Y^+, Y^-] \in \mathcal{S}_{\theta}.$$

The frequency-support conditions follow directly from the definition of Γ . As a result, it remains to prove that

$$\|\Gamma[B, X^+, X^-, Y^+, Y^-]\|_{\mathcal{S}} \leq \theta. \quad (8.49)$$

By symmetry in the u and v -variables, it suffices to prove the required bounds in (8.49) for the Γ_B , $(\Gamma_X)^+$, and $(\Gamma_Y)^+$ -components.

Step 2.i: Self-mapping bound for the Γ_B -component. Using the bilinear estimate (Proposition 2.5), the composition estimate (Lemma 2.9), and Definition 3.12, we have that

$$\|(\Gamma_B)_{M,N}\|_{C_u^s C_v^s} \lesssim \|\mathbb{I}^\diamond(\phi)\|_{C_u^s C_v^s} \|A_M^+\|_{C_u^s C_v^s} \|A_N^-\|_{C_u^s C_v^s} \lesssim \theta^2.$$

Since $\theta > 0$ is sufficiently small, we obtain that $\|(\Gamma_B)_{M,N}\|_{C_u^s C_v^s} \leq 5\theta$, which is the desired estimate.

Step 2.ii: Self-mapping bound for the $(\Gamma_X)^+$ -component. It suffices to estimate $\mathcal{X}_{M,N,mn}^{+,a}$ in $C_u^s C_v^s$. Using the bilinear estimate (Proposition 2.5), the bilinear estimate for families (Corollary 2.7), and the composition estimate (Lemma 2.9), we have that

$$\begin{aligned} \|(\Gamma_X)_{M,m}^{+,a}\|_{C_u^s C_v^s} &\leq \|\mathcal{X}_{M,N,mn}^{+,a}\|_{C_u^s C_v^s} \\ &\leq \|(8.13)\|_{C_u^s C_v^s} + \|(8.14)\|_{C_u^s C_v^s} + \|(8.15)\|_{C_u^s C_v^s} \\ &\lesssim \|\mathbb{I}^\diamond(\phi)\|_{C_u^s C_v^s} \|A_M^+\|_{C_u^s C_v^s} \|A_N^-\|_{C_u^s C_v^s} \\ &\quad + \sup_L \|\mathbb{I}^\diamond(\phi)\|_{C_u^s C_v^s} \|A_M^+\|_{C_u^s C_v^s} \|\phi_L^+\|_{C_u^s} \|X_{L,N}^+\|_{C_u^s C_v^s} \\ &\quad + \sup_K \|\mathbb{I}^\diamond(\phi)\|_{C_u^s C_v^s} \|A_M^+\|_{C_u^s C_v^s} \|B_{K,N}\|_{C_u^s C_v^s} \|\phi_K^+\|_{C_u^s} \\ &\lesssim \theta^2. \end{aligned}$$

Since $\theta > 0$ is sufficiently small, this implies the desired estimate.

Step 2.iii: Self-mapping bound for the $(\Gamma_Y)^+$ -component. In this step, we prove that

$$\|(\Gamma_Y)^+ - \theta \text{Id}\|_{C_u^s C_v^r} \lesssim \theta^2,$$

which implies the desired estimate. Using the definition of $(\Gamma_Y)^+$, it suffices to prove that

$$\left\| \sum_{N \leq M^{1-\delta}} \text{Com}_{\chi, \otimes, I}^v(\mathcal{X}_{M,N,mn}^{+,a}, \phi_N^{-,n}) \right\|_{C_u^s C_v^r} \lesssim \theta^2, \quad (8.50)$$

$$\|\mathcal{Y}_{M,m}^{+,a}\|_{C_u^s C_v^{r-1}} \lesssim \theta^3. \quad (8.51)$$

In Step 2.ii, we have already estimated $\mathcal{X}_{M,N}^+$ in $C_u^s C_v^s$. Then, the estimate (8.50) directly follows from Lemma 2.13. To prove the estimate (8.51), we further distinguish between the four components $\mathcal{Y}^{(1)}$, $\mathcal{Y}^{(2)}$, $\mathcal{Y}^{(3)}$, and $\mathcal{Y}^{(4)}$.

Estimate of $\mathcal{Y}^{(1)}$: Using Lemma 2.13 for \odot^v , the bilinear estimate for families (Corollary 2.7), and the multiplication estimate (Corollary 2.37), it remains to control

$$A_{M,m}^{+,i} \odot^v \partial_v \phi_N^{-,n}, \quad \mathbb{I}_{ij}^{\diamond,a}(\phi) \odot^v \partial_v \phi_N^{-,n}, \quad (8.52)$$

$$A_{N,n}^{-,j} \odot^v \partial_v \phi_N^{-,n}, \quad \text{and} \quad B_{K,N,kn}^j \odot^v \partial_v \phi_N^{-,n} \quad (8.53)$$

in $C_u^s C_v^{r-1}$. The two terms in (8.52) are controlled by Lemma 8.5. The two terms in (8.53) vanish due to the frequency-support conditions on $A_{N,n}^{-,j}$ and $B_{K,N,kn}^j$.

Estimate of $\mathcal{Y}^{(2)}$: This term is the content of Lemma 8.6.

Estimate of $\mathcal{Y}^{(3)}$: This term can be bounded directly through Lemma 2.13.

Estimate of $\mathcal{Y}^{(4)}$: The remaining “easy” terms can be bounded using the bilinear estimate (Proposition 2.5) and Lemma 4.5. As a result, we omit the details.

This completes the proof of the second estimate (8.51) and, therefore, the proof of the self-mapping estimate (8.49).

Step 3: The contraction estimate and continuous dependence on the data. In Step 2, we proved that Γ maps \mathcal{S}_θ back into itself. In order to utilize the contraction mapping theorem, it remains to prove that Γ is a contraction on \mathcal{S}_θ , which would follow from the estimate

$$\begin{aligned} & \|\Gamma[B, X^+, X^-, Y^+, Y^-] - \Gamma[B', (X')^+, (X')^-, (Y')^+, (Y')^-]\|_{\mathcal{S}} \\ & \lesssim \theta \|B, X^+, X^-, Y^+, Y^- - B', (X')^+, (X')^-, (Y')^+, (Y')^-\|_{\mathcal{S}}. \end{aligned} \quad (8.54)$$

Compared to our previous estimate (8.49), the only new difficulties in proving (8.54) are notational, since Γ contains a large number of terms. The individual estimates, however, are the exact same estimates as used in the proof of (8.49). As a result, we omit the (extremely tedious but standard) details.

As is common for contraction mapping arguments, our estimates also yield the Lipschitz-continuous dependence on the data, which is given by $(\phi_0(0), \phi^+, \phi^-, \psi) \in \mathcal{M} \times \mathcal{D}^s \times C_u^r C_v^r$. This can be shown by further generalizing the contraction estimates (8.54). After reflecting the dependence on $(\phi_0(0), \phi^+, \phi^-, \psi)$ in our notation, the desired estimate reads

$$\begin{aligned} & \|\Gamma[\phi_0(0), \phi^+, \phi^-, \psi, B, X^+, X^-, Y^+, Y^-] \\ & - \Gamma[\phi'_0(0), (\phi')^+, (\phi')^-, \psi', B', (X')^+, (X')^-, (Y')^+, (Y')^-]\|_{\mathcal{S}} \\ & \lesssim \theta \left(\|\phi_0(0) - \phi'_0(0)\|_{\mathbb{R}^D} + \|(\phi^+, \phi^-); ((\phi')^+, (\phi')^-)\|_{\mathcal{D}^s} + \|\psi - \psi'\|_{C_u^r C_v^r} \right. \\ & \quad \left. + \|(B, X^+, X^-, Y^+, Y^-) - (B', (X')^+, (X')^-, (Y')^+, (Y')^-)\|_{\mathcal{S}} \right). \end{aligned} \quad (8.55)$$

As for the contraction estimate (8.54), its generalization (8.55) can be proven using the estimates leading to (8.49). We therefore omit the (extremely tedious but standard) details. \square

9. Local Well-Posedness

In the final section of the article, we present the proof of Proposition 3.16 and Theorem 1.4.

Proof of Proposition 3.16. Throughout the proof, we assume that $\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta$ and restrict our discussion to $\psi \in C_u^r C_v^r$ satisfying $\|\psi\|_{C_u^r C_v^r} \leq \theta$. We now split the argument into two steps. In the first step, we simplify the forced wave maps equation (3.45). In the second step, we solve the simplified version using a contraction mapping argument.

Step 1: The simplified forced wave maps equation. Due to Proposition 3.14, there exist modulations $A^+[\psi]$, $A^-[\psi]$, and $B[\psi]$ satisfying the modulation equations from Definition 3.13. We recall from (3.44) that

$$\begin{aligned} \phi^k[\psi](u, v) &= \sum_M A_{M,m}^{+,k}[\psi](u, v) \phi_M^{+,m}(u) + \sum_N A_{N,n}^{-,k}[\psi](u, v) \phi_N^{-,n}(v) \\ &+ \sum_{\substack{M,N: \\ M \sim_\delta N}} B_{M,N,mn}^k[\psi](u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v) + \psi^k(u, v). \end{aligned} \quad (9.1)$$

Using the modulation equations and Proposition 7.1, we have that

$$\begin{aligned} & \theta(\phi^{+,k} + \phi^{-,k}) - \chi^+ \chi^- \text{Duh} \left[\mathbb{I}_{ij}^k(\phi[\psi]) \partial_u \phi^i[\psi] \partial_v \phi^j[\psi] \right] \\ &= \sum_{\substack{M,N: \\ M \sim_\delta N}} B_{M,N,mn}^k[\psi](u, v) \phi_M^{+,m}(u) \phi_N^{-,n}(v) + \sum_M A_{M,m}^{+,k}[\psi](u, v) \phi_M^{+,m}(u) \\ & \quad + \sum_N A_{N,n}^{-,k}[\psi](u, v) \phi_N^{-,n}(v) + \mathcal{R}^k[\phi_0(0), \phi^\pm, A^\pm[\psi], B[\psi], \psi]. \end{aligned} \quad (9.2)$$

Here, the remainder $\mathcal{R} = \mathcal{R}[\phi_0(0), \phi^\pm, A^\pm, B, \psi]$ is as in the statement of Proposition 7.1. By combining (9.1) and (9.2), all terms but the remainder cancel, which yields

$$\begin{aligned} & \theta(\phi^{+,k} + \phi^{-,k}) - \chi^+ \chi^- \text{Duh} \left[\mathbb{I}_{ij}^k(\phi[\psi]) \partial_u \phi^i[\psi] \partial_v \phi^j[\psi] \right] - (\phi^k[\psi] - \psi^k) \\ &= \mathcal{R}^k[\phi_0(0), \phi^\pm, A^\pm[\psi], B[\psi], \psi]. \end{aligned}$$

As a result, the forced wave maps equation (3.45) reads

$$\psi^k = \mathcal{R}^k[\phi_0(0), \phi^\pm, A^\pm[\psi], B[\psi], \psi]. \quad (9.3)$$

We call (9.3) the simplified forced wave maps equation, which we now solve using a contraction mapping argument.

Step 2: Solving the simplified forced wave maps equation. We define the solution set

$$\mathcal{S}_\theta = \{\psi \in C_u^r C_v^r : \|\psi\|_{C_u^r C_v^r} \leq \theta\},$$

which is equipped with the $C_u^r C_v^r$ -norm. To simplify the notation, we write

$$\Gamma[\phi_0(0), \phi^\pm, \psi] := \mathcal{R}[\phi_0(0), \phi^\pm, A^\pm[\psi], B[\psi], \psi]. \quad (9.4)$$

We now let $\psi, \tilde{\psi} \in \mathcal{S}_\theta$ and assume that $(\phi^+, \phi^-), (\tilde{\phi}^+, \tilde{\phi}^-) \in \mathcal{D}^s$ satisfy

$$\|(\phi^+, \phi^-)\|_{\mathcal{D}^s} \leq \theta \quad \text{and} \quad \|(\tilde{\phi}^+, \tilde{\phi}^-)\|_{\mathcal{D}^s} \leq \theta.$$

In order to prove that Γ is a contraction on \mathcal{S}_θ and that the resulting fixed-point depends continuously on the data, it suffices to prove the two estimates

$$\|\Gamma[\phi_0(0), \phi^\pm, \psi]\|_{C_u^r C_v^r} \leq \theta \quad (9.5)$$

and

$$\begin{aligned} & \|\Gamma[\phi_0(0), \phi^\pm, \psi] - \Gamma[\tilde{\phi}_0(0), \tilde{\phi}^\pm, \tilde{\psi}]\|_{C_u^r C_v^r} \\ & \lesssim \theta (\|\phi_0(0) - \tilde{\phi}_0(0)\|_{\mathbb{R}^D} + \|(\phi^+, \phi^-); (\tilde{\phi}^+, \tilde{\phi}^-)\|_{\mathcal{D}^s} + \|\psi - \tilde{\psi}\|_{C_u^r C_v^r}). \end{aligned} \quad (9.6)$$

The first inequality (9.5) follows directly from Proposition 3.14 and Proposition 7.1. Indeed, we obtain from Proposition 3.14 that $A^+[\psi]$, $A^-[\psi]$, and $B[\psi]$ satisfy Hypothesis 3.8. Then, it follows from Proposition 7.1 that

$$\|\Gamma[\phi_0(0), \phi^\pm, \psi]\|_{C_u^r C_v^r} = \|\mathcal{R}[\phi_0(0), \phi^\pm, A^\pm[\psi], B[\psi], \psi]\|_{C_u^r C_v^r} \lesssim \theta^2 \leq \theta.$$

Compared to the first inequality (9.5), the only new difficulties in proving the second inequality (9.6) are notational. The reason is that, once we insert the expression for \mathcal{R} implicit from (7.5), the left-hand side of (9.6) contains numerous terms. The individual estimates, however, are the exact same estimates as those leading to the proof of (9.5). As a result, we omit the (extremely tedious but standard) details. \square

We now prove the local well-posedness of the wave maps equation (WM) for small data in \mathcal{D}^s . As we show below, the main theorem then follows from a scaling argument.

Proposition 9.1 (Local well-posedness for small data). *Let $(\phi_0^\varepsilon, \phi_1^\varepsilon)_\varepsilon \subseteq C^\infty(\mathbb{R} \rightarrow T\mathcal{M})$ and assume that the following three conditions are satisfied:*

- (i) *The initial positions $\phi_0^\varepsilon(0)$ converges in \mathcal{M} .*
- (ii) *The shifted linear waves $\phi^{\diamond,+, \varepsilon}$ and $\phi^{\diamond,-, \varepsilon}$ converge in \mathcal{D}^s .*
- (iii) *The shifted linear waves $\phi^{\diamond,+, \varepsilon}$ and $\phi^{\diamond,-, \varepsilon}$ are small in \mathcal{D}^s , i.e.,*

$$\sup_{\varepsilon > 0} \|(\phi^{\diamond,+, \varepsilon}, \phi^{\diamond,-, \varepsilon})\|_{\mathcal{D}^s} \leq \theta.$$

Then, the smooth global solutions ϕ^ε of (WM) with initial data $(\phi_0^\varepsilon, \phi_1^\varepsilon)$ converge in

$$(C_t^0 C_x^s \cap C_t^1 C_x^{s-1})([-1, 1] \times [-1, 1] \rightarrow \mathcal{M}). \quad (9.7)$$

Proof of Proposition 9.1: We define the shifted wave map $\phi^{\diamond, \varepsilon} : \mathbb{R}_{u,v}^{1+1} \rightarrow \mathbb{R}^D$ by

$$\phi^{\diamond, \varepsilon}(u, v) = \phi^\varepsilon(u, v) - \phi_0^\varepsilon(0).$$

Since ϕ^ε is a global smooth solution of (WM), the shifted wave map is a solution of

$$\phi^{\diamond, \varepsilon, k} = \phi^{\diamond,+, \varepsilon, k} + \phi^{\diamond,-, \varepsilon, k} - \text{Duh}[\Pi_{ij}^{\diamond, k}(\phi^{\diamond, \varepsilon}) \partial_u \phi^{\diamond, \varepsilon, i} \partial_v \phi^{\diamond, \varepsilon, j}]. \quad (9.8)$$

In addition to (9.8), we consider the localized Duhamel integral formulation

$$\tilde{\phi}^{\diamond, \varepsilon, k} = \phi^{\diamond,+, \varepsilon, k} + \phi^{\diamond,-, \varepsilon, k} - \chi^+ \chi^- \text{Duh}[\Pi_{ij}^{\diamond, k}(\tilde{\phi}^{\diamond, \varepsilon}) \partial_u \tilde{\phi}^{\diamond, \varepsilon, i} \partial_v \tilde{\phi}^{\diamond, \varepsilon, j}]. \quad (9.9)$$

Using the assumptions, Proposition 3.14, and Proposition 3.16, we obtain the decomposition

$$\begin{aligned} \tilde{\phi}^{\diamond, \varepsilon, k} &= \sum_M A_{M,m}^{+, \varepsilon, k} \phi_M^{\diamond,+, \varepsilon, m} + \sum_N A_{N,n}^{-, \varepsilon, k} \phi_N^{\diamond,-, \varepsilon, n} \\ &\quad + \sum_{M \sim_\delta N} B_{M,N,mn}^\varepsilon \phi_M^{\diamond,+, \varepsilon, m} \phi_N^{\diamond,-, \varepsilon, n} + \psi^{\varepsilon, k}, \end{aligned} \quad (9.10)$$

where the modulations and nonlinear remainder satisfy the following properties:

- (i) The modulations $A^{+, \varepsilon}$, $A^{-, \varepsilon}$, and B^ε satisfy Condition 3.5.
- (ii) The modulations $A^{+, \varepsilon}$, $A^{-, \varepsilon}$, and B^ε converge with respect to $\|\cdot\|_{Mod}$ from Definition 3.6.
- (iii) The nonlinear remainder ψ^ε converges in $C_u^r C_v^r$.
- (iv) The modulations and nonlinear remainder are small, i.e.,

$$\sup_{\varepsilon > 0} \left(\frac{1}{2} \sup_M \|A_M^{+, \varepsilon}\|_{Mod_M^+} + \frac{1}{2} \sup_N \|A_N^{-, \varepsilon}\|_{Mod_N^-} + \sup_{M \sim_\delta N} \|B_{M,N}^\varepsilon\|_{C_u^s C_v^s} + \|\psi^\varepsilon\|_{C_u^r C_v^r} \right) \leq \theta.$$

For every fixed $\varepsilon > 0$, it follows from the smoothness of the shifted linear waves, the decomposition (9.10), and the properties of $A^{+, \varepsilon}$, $A^{-, \varepsilon}$, B^ε , and ψ^ε that $\phi^{\diamond, \varepsilon, k} \in C_u^r C_v^r$. From the deterministic theory (Lemma 2.27), it follows that

$$\tilde{\phi}^{\diamond, \varepsilon}(u, v) = \phi^{\diamond, \varepsilon}(u, v) \quad \text{for all } u, v \in [-2, 2]. \quad (9.11)$$

In order to prove the proposition, it therefore suffices to prove the convergence of $\tilde{\phi}^{\diamond, \varepsilon}$ in (9.7). However, this follows directly from Lemma 3.11. \square

Equipped with Proposition 9.1, we are now ready to prove the main theorem.

Proof of Theorem 1.4. The proof consists of two steps: In the first step, we perform a series of changes of variables which lead to small initial data. In the second step, we apply Proposition 9.1 and conclude the theorem.

Step 1: Change of variables. We first recall that ϕ^ε solves the wave maps equation with initial data $(B^\varepsilon, V^\varepsilon)$, i.e.,

$$\begin{cases} \partial_\mu \partial^\mu \phi^{\varepsilon,k} = -\mathbb{I}_{ij}^k(\phi^\varepsilon) \partial_\mu \phi^\varepsilon \partial^\mu \phi^\varepsilon, \\ \phi^\varepsilon(0, x) = B^\varepsilon(x), \quad \partial_t \phi^\varepsilon(0, x) = V^\varepsilon(x). \end{cases} \quad (9.12)$$

We note that since $B^\varepsilon: \mathbb{R} \rightarrow \mathcal{M} \subseteq \mathbb{R}^D$ and $V^\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^D$ are smooth and satisfy $V^\varepsilon(x) \in T_{B^\varepsilon(x)}\mathcal{M}$, (9.12) has a unique global smooth solution. We now perform a re-scaling, spatial translation, and localization.

Step 1.1: Re-scaling and spatial translation. Let $x_0 \in \mathbb{R}$ be arbitrary. We define the re-scaled and translated solution $\phi_{\tau, x_0}^\varepsilon$, Brownian path $B_{\tau, x_0}^\varepsilon$, and velocity $V_{\tau, x_0}^\varepsilon$ by

$$\phi_{\tau, x_0}^\varepsilon(t, x) := \phi^\varepsilon(\tau t, \tau x + x_0), \quad B_{\tau, x_0}^\varepsilon(x) = B^\varepsilon(\tau x + x_0), \quad \text{and} \quad V_{\tau, x_0}^\varepsilon(x) = \tau V^\varepsilon(\tau x + x_0).$$

Due to the scale and translation-invariance of the wave maps equation, $\phi_{\tau, x_0}^\varepsilon$ solves

$$\begin{cases} \partial_\mu \partial^\mu \phi_{\tau, x_0}^\varepsilon = -\mathbb{I}_{ij}^k(\phi_{\tau, x_0}^\varepsilon) \partial_\mu \phi_{\tau, x_0}^\varepsilon \partial^\mu \phi_{\tau, x_0}^\varepsilon, \\ \phi_{\tau, x_0}^\varepsilon(0, x) = B_{\tau, x_0}^\varepsilon(x), \quad \partial_t \phi_{\tau, x_0}^\varepsilon(0, x) = V_{\tau, x_0}^\varepsilon(x). \end{cases} \quad (9.13)$$

Step 1.2: Localization. We let $B_{\tau, x_0, \text{loc}}^\varepsilon$ and $V_{\tau, x_0, \text{loc}}^\varepsilon$ be the localized versions $B_{\tau, x_0}^\varepsilon$ and $V_{\tau, x_0}^\varepsilon$, which were defined in (2.96) and (2.134). We then define $\phi_{\tau, x_0, \text{loc}}^\varepsilon$ as the solution to the wave maps equation with initial data $(B_{\tau, x_0, \text{loc}}^\varepsilon, V_{\tau, x_0, \text{loc}}^\varepsilon)$. Due to finite speed of propagation, it holds that

$$\phi_{\tau, x_0, \text{loc}}^\varepsilon(t, x) = \phi_{\tau, x_0}^\varepsilon(t, x) \quad \text{for all } (t, x) \in [-1, 1]^2. \quad (9.14)$$

Step 1.3: Shifts. Similar as in (3.2), we define the shifted linear waves by

$$\phi_{\tau, x_0, \text{loc}}^{\diamond, \pm, \varepsilon}(x) := \frac{1}{2\theta} \left(B_{\tau, x_0, \text{loc}}^\varepsilon(x) - B_{\tau, x_0, \text{loc}}^\varepsilon(0) \mp \int_0^x dy V_{\tau, x_0, \text{loc}}^\varepsilon(y) \right).$$

By Proposition 2.32, Proposition 2.33, and Proposition 2.35, there exists an event $\mathcal{E}(\tau, x_0)$ satisfying

$$\mathbb{P}(\Omega \setminus \mathcal{E}(\tau, x_0)) \leq C \exp(-c\tau^{-c}) \quad (9.15)$$

for some absolute constant $c > 0$ and a constant $C = C(\mathcal{M})$ depending on the embedding $\mathcal{M} \hookrightarrow \mathbb{R}^D$, and such that, on the event $\mathcal{E}(\tau, x_0)$, the following properties hold:

- (i) The initial position $B_{\tau, x_0, \text{loc}}^\varepsilon(0)$ converges in \mathcal{M} .
- (ii) The shifted linear waves $\phi_{\tau, x_0, \text{loc}}^{\diamond, +, \varepsilon}(x)$ and $\phi_{\tau, x_0, \text{loc}}^{\diamond, -, \varepsilon}(x)$ converge in \mathcal{D}^s .
- (iii) The shifted linear waves $\phi_{\tau, x_0, \text{loc}}^{\diamond, +, \varepsilon}(x)$ and $\phi_{\tau, x_0, \text{loc}}^{\diamond, -, \varepsilon}(x)$ are small in \mathcal{D}^s , i.e.,

$$\sup_{\varepsilon > 0} \left\| (\phi_{\tau, x_0, \text{loc}}^{\diamond, +, \varepsilon}, \phi_{\tau, x_0, \text{loc}}^{\diamond, -, \varepsilon}) \right\|_{\mathcal{D}^s} \leq \theta.$$

Step 2: Conclusion. We proceed on the event $\mathcal{E}(\tau, x_0)$. Using the properties in Step 1.3 and Proposition 9.1, it follows that

$$\phi_{\tau, x_0, \text{loc}}^{\varepsilon} \text{ converges in } (C_t^0 C_x^s \cap C_t^1 C_x^{s-1})([-1, 1]^2 \rightarrow \mathcal{M}).$$

Due to finite speed of propagation (9.14), we obtain that

$$\phi_{\tau, x_0}^{\varepsilon} \text{ converges in } (C_t^0 C_x^s \cap C_t^1 C_x^{s-1})([-1, 1]^2 \rightarrow \mathcal{M}).$$

By undoing the scaling and spatial translation, it follows that

$$\phi^{\varepsilon} \text{ converges in } (C_t^0 C_x^s \cap C_t^1 C_x^{s-1})([-\tau, \tau] \times [x_0 - \tau, x_0 + \tau] \rightarrow \mathcal{M}). \quad (9.16)$$

We now define the event $\mathcal{E}(\tau, R)$ in the statement of the theorem by

$$\mathcal{E}(\tau, R) := \bigcap_{n=-\lceil 4R/\tau \rceil}^{\lceil 4R/\tau \rceil} \mathcal{E}(\tau, \tau n),$$

where $\lceil \cdot \rceil$ is the ceiling function. The estimate on the probability of $\mathcal{E}(\tau, R)$ follows from (9.15) and a union bound. The convergence statement on $[-\tau, \tau] \times [-R, R]$ follows directly from (9.16) and a partition of unity in space. \square

Acknowledgements The authors thank Rishabh Gvalani, Florian Kunich, Stephan Luckhaus, Felix Otto, Igor Rodnianski, Angela Stevens, Terence Tao, and Markus Tempelmayr for helpful and interesting discussions. The authors also thank the anonymous referees for valuable comments and suggestions. B.B. thanks the MPI for Mathematics in the Sciences for support during a visit in the summer of 2021. The three authors thank ICERM, which is supported by NSF grant DMS-1929284, for support during the semester program on Hamiltonian Methods in Dispersive and Wave Evolution Equations. B.B. was partially supported by the NSF under Grant No. DMS-1926686. J.L. was partially supported by NSF grant DMS-1954707. G.S. was partially supported by DMS-1764403, DMS-2052651 and the Simons Foundation.

Data Availability This article does not have any external supporting data.

Declarations

Conflict of interest The authors do not have any competing interests to declare.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

A. Appendix A. Deterministic Ill-Posedness

We now prove the mild ill-posedness statement in Theorem 1.2, which concerns the unboundedness of the first Picard iterate.

Proof of Theorem 1.2(ii). We only treat the case \mathbb{S}^2 , since the argument easily generalizes to \mathbb{S}^{D-1} with $D \geq 3$. We split the argument into two steps. In the first step, we present several reductions which simplify the first Picard iterate. In the second step, we construct an explicit sequence of functions for which the first Picard iterate diverges.

Step 1: Reductions. We let $\phi_0, \phi_1 \in C_b^\infty(\mathbb{R} \rightarrow \mathbb{R}^3)$ satisfy

$$\|\phi_0(x)\|_2^2 = 1 \quad \text{and} \quad \langle \phi_0(x), \phi_1(x) \rangle = 0 \quad \forall x \in \mathbb{R}. \quad (\text{A.1})$$

Here, $\|\cdot\|_2$ refers to the Euclidean norm on \mathbb{R}^3 . Then, the right and left-moving linear waves are given by

$$\phi^\pm(x) = \frac{1}{2} \left(\phi_0(x) \mp \int_0^x dy \phi_1(y) \right).$$

We recall that the second fundamental form of the sphere \mathbb{S}^2 is given by

$$\mathbb{I}_{ij}^k(\phi) = \delta_{ij} \phi^k.$$

As a result, the first Picard iterate of the wave maps equation (WM) is given by

$$\mathcal{P}(t, x) = - \int_{x-t}^{x+t} dv' \int_{x-t}^{v'} du' (\phi^+(u') + \phi^-(v')) \langle (\partial_u \phi^+)(u'), (\partial_v \phi^-)(v') \rangle.$$

We now fix any time $t > 0$. Furthermore, we let χ be our previous nonnegative, smooth cut-off function, which satisfies $\text{supp}(\chi) \subseteq (-3, 3)$, and let $0 < \epsilon \leq t/100$. For any $r \in \mathbb{R}$, it then holds that

$$\left\| \int_{-\infty}^{\infty} \chi(x/\epsilon) \mathcal{P}(t, x) dx \right\|_2 \lesssim_{\chi, \epsilon} \|\mathcal{P}(t, \cdot)\|_{C^r}.$$

In order to prove the unboundedness of the first Picard iterate on $C^r \times C^{r-1}$ for any $r \leq 1/2$, it therefore suffices to prove that

$$\sup_{\|\phi_0\|_{C^{1/2}} \leq 1} \sup_{\|\phi_1\|_{C^{-1/2}} \leq 1} \left\| \int_{-\infty}^{\infty} \chi(x/\epsilon) \mathcal{P}(t, x) dx \right\|_2 = \infty. \quad (\text{A.2})$$

We now choose $\phi_0(x) = e_3 \in \mathbb{R}^3$ and choose $\phi_1(x) = \psi'(x)$ for $\psi(x) \in C_c^\infty((-1, 1) \rightarrow \mathbb{R}^3)$. In particular, it holds that $\|\psi\|_{C^{1/2}} \lesssim \|\phi_1\|_{C^{-1/2}}$. Due to the geometric constraint in (A.1), ψ has to satisfy $\langle e_3, \psi \rangle = 0$. Using our assumptions, the first Picard iterate takes the form

$$\mathcal{P}(t, x) = -\frac{1}{8} \int_{x-t}^{x+t} dv' \int_{x-t}^{v'} du' (2e_3 - \psi(u') + \psi(v')) \langle \psi'(u'), \psi'(v') \rangle.$$

In order to further simplify $\mathcal{P}(t, x)$, we assume that $\text{supp}(\psi) \subseteq (-2\epsilon, 2\epsilon)$. In particular, it holds that $\psi(x+t) = \psi(x-t) = 0$ for all $x \in (-2\epsilon, 2\epsilon)$. Then, a direct computation yields that

$$\begin{aligned} & \int_{x-t}^{x+t} dv' \int_{x-t}^{v'} du' 2e_3 \langle \psi'(u'), \psi'(v') \rangle = 0, \\ & - \int_{x-t}^{x+t} dv' \int_{x-t}^{v'} du' \psi(u') \langle \psi'(u'), \psi'(v') \rangle = -\frac{1}{2} \int_{x-t}^{x+t} dy \psi'(y) \|\psi(y)\|_2^2, \\ & \int_{x-t}^{x+t} dv' \int_{x-t}^{v'} du' \psi(v') \langle \psi'(u'), \psi'(v') \rangle = -\frac{1}{2} \int_{x-t}^{x+t} dy \psi'(y) \|\psi(y)\|_2^2. \end{aligned}$$

As a result, we obtain that

$$\mathcal{P}(t, x) = \frac{1}{8} \int_{x-t}^{x+t} dy \, \psi'(y) \|\psi(y)\|_2^2.$$

We now write $\psi(y) = \psi^1(y)e_1 + \psi^2(y)e_2$, which satisfies the constraint $\langle \psi, e_3 \rangle = 0$. Then, the first coordinate of $\mathcal{P}(t, x)$ is given by

$$\begin{aligned} \langle \mathcal{P}(t, x), e_1 \rangle &= \frac{1}{8} \int_{x-t}^{x+t} dy \, (\psi^1)'(y) ((\psi^1)^2(y) + (\psi^2)^2(y)) \\ &= \frac{1}{24} (\psi^1)^3(y) \Big|_{y=x-t}^{x+t} + \frac{1}{8} \int_{x-t}^{x+t} dy \, (\psi^1)'(y) (\psi^2)^2(y) \\ &= \frac{1}{8} \int_{x-t}^{x+t} dy \, (\psi^1)'(y) (\psi^2)^2(y). \end{aligned}$$

Thus, it suffices to prove that

$$\sup_{\substack{\psi^1, \psi^2: \\ \text{supp}(\psi^j) \subseteq (-2\epsilon, 2\epsilon) \\ \|\psi^j\|_{C^{1/2}} \leq 1}} \left| \int_{-\infty}^{\infty} dx \, \chi(x/\epsilon) \int_{x-t}^{x+t} dy \, (\psi^1)'(y) (\psi^2)^2(y) \right| = \infty. \quad (\text{A.3})$$

Step 2: Proof of (A.3). The main idea in the proof of (A.3) is to create a severe high \times high \times low-interaction in the integrand. In order to cover the endpoint $C^{1/2}$, however, we need to be careful and work at multiple scales.

We let $\kappa, \kappa_0 \in \mathbb{N}$ be arbitrary and define the set of frequencies

$$F_{\kappa, \kappa_0} = \{2^{10k} : \kappa_0 \leq k \leq \kappa\}.$$

Then, we define

$$\psi^1(y) = \chi(y/\epsilon) \sum_{n \in F_{\kappa, \kappa_0}} n^{-1/2} \sin(ny), \quad (\text{A.4})$$

$$\psi^2(y) = \chi(y/\epsilon) \left(\sin(y) + \sum_{n \in F_{\kappa, \kappa_0}} n^{-1/2} \sin((n-1)y) \right). \quad (\text{A.5})$$

Since the frequencies in F_{κ, κ_0} are well-separated, it holds that

$$\|\psi^1\|_{C^{1/2}}, \|\psi^2\|_{C^{1/2}} \lesssim_{\epsilon} 1,$$

where the implicit constant is uniform in κ and κ_0 . We now claim that

$$\begin{aligned} &\left| \int_{x-t}^{x+t} dy \, (\psi^1)'(y) (\psi^2)^2(y) + \frac{1}{2} (\kappa - \kappa_0) \int_{x-t}^{x+t} dy \, \chi(y/\epsilon)^3 (1 - \cos(2y)) \right| \\ &\lesssim_{\epsilon} 2^{-5\kappa_0} (\kappa - \kappa_0) + 1. \end{aligned} \quad (\text{A.6})$$

Before proving (A.6), we first show that (A.6) implies (A.3). By integrating (A.6) against $\chi(x/\epsilon)$, using that $1 - \cos(2y) \geq 0$, and using that $0 < \epsilon \leq t/10$, we obtain

$$\left| \int_{-\infty}^{\infty} dx \, \chi(x/\epsilon) \int_{x-t}^{x+t} dy \, (\psi^1)'(y) (\psi^2)^2(y) \right|$$

$$\begin{aligned} &\geq \frac{1}{2}(\kappa - \kappa_0) \left| \int_{-\infty}^{\infty} dx \chi(x/\epsilon) \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 (1 - \cos(2y)) \right| - C_\epsilon (2^{-5\kappa_0}(\kappa - \kappa_0) + 1) \\ &\geq c_\epsilon(\kappa - \kappa_0) - C_\epsilon (2^{-5\kappa_0}(\kappa - \kappa_0) + 1), \end{aligned}$$

where $C_\epsilon > 0$ and $c_\epsilon > 0$ are sufficiently large and small constants, respectively. We now obtain the desired conclusion (A.3) by first choosing a parameter $\kappa_0 = \kappa_0(\epsilon)$ such that $C_\epsilon 2^{-5\kappa_0}$ is smaller than c_ϵ and then letting $\kappa \rightarrow \infty$. Thus, it now only remains to prove the claim (A.6). By inserting (A.4) and (A.5) into the integrand, we obtain that

$$\begin{aligned} &\int_{x-t}^{x+t} dy (\psi^2)^2(y) (\psi^1)'(y) \\ &= 2 \sum_{m,n \in F_{\kappa, \kappa_0}} m^{-1/2} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin(y) \sin((m-1)y) \cos(ny) \end{aligned} \quad (\text{A.7})$$

$$+ \sum_{n \in F_{\kappa, \kappa_0}} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin^2(y) \cos(ny) \quad (\text{A.8})$$

$$+ \sum_{\ell, m, n \in F_{\kappa, \kappa_0}} \ell^{-1/2} m^{-1/2} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin((\ell-1)y) \sin((m-1)y) \cos(ny) \quad (\text{A.9})$$

$$+ \mathcal{O}_\epsilon(1),$$

where we have already estimated terms in which the derivative hits the cut-off $\chi(y/\epsilon)$. We start by analyzing the main term (A.7). First, we treat the contribution of the diagonal case $m = n$. Using trigonometric identities, we have that

$$\sin(y) \sin((n-1)y) \cos(ny) = \frac{1}{4} \left(-1 + \cos(2y) + \cos((2n-2)y) - \cos(2ny) \right).$$

Using integration by parts, this yields

$$\begin{aligned} &2 \sum_{\substack{m, n \in F_{\kappa, \kappa_0} \\ m=n}} m^{-1/2} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin(y) \sin((m-1)y) \cos(ny) \\ &= -\frac{1}{2} \sum_{n \in F_{\kappa, \kappa_0}} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 (1 - \cos(2y)) + \mathcal{O}_\epsilon \left(\sum_{n \in F_{\kappa, \kappa_0}} n^{-1} \right) \\ &= -\frac{1}{2}(\kappa - \kappa_0) \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 (1 - \cos(2y)) + \mathcal{O}_\epsilon(1), \end{aligned}$$

which is the main term in (A.6). In the non-diagonal case $m \neq n$, we use that the frequencies in F_{κ, κ_0} are well-separated. Together with integration by parts, this yields

$$\begin{aligned} &\left| \sum_{\substack{m, n \in F_{\kappa, \kappa_0} \\ m \neq n}} m^{-1/2} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin(y) \sin((m-1)y) \cos(ny) \right| \\ &\lesssim_\epsilon \sum_{m, n \in F_{\kappa, \kappa_0}} m^{-1/2} n^{1/2} \max(m, n)^{-1} \lesssim 1. \end{aligned}$$

We now estimate the first error term (A.8). Using integration by parts, it follows that

$$\left| \sum_{n \in F_{\kappa, \kappa_0}} n^{1/2} \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin^2(y) \cos(ny) \right| \lesssim_{\epsilon} \sum_{n \in F_{\kappa, \kappa_0}} n^{-\frac{1}{2}} \lesssim 1.$$

It remains to estimate the third error term (A.9). To this end, we distinguish two cases. In the case $\max(\ell, m, n) > \text{med}(\ell, m, n)$, we use that the frequencies in F_{κ, κ_0} are well-separated, which implies that

$$\begin{aligned} & \left| \sum_{\substack{\ell, m, n \in F_{\kappa, \kappa_0} : \\ \max(\ell, m, n) > \text{med}(\ell, m, n)}} \ell^{-1/2} m^{-1/2} n^{1/2} \right. \\ & \quad \left. \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin((\ell-1)y) \sin((m-1)y) \cos(ny) \right| \\ & \lesssim_{\epsilon} \sum_{\ell, m, n \in F_{\kappa, \kappa_0}} \ell^{-1/2} m^{-1/2} n^{1/2} \max(\ell, m, n)^{-1} \lesssim 1. \end{aligned}$$

In the case $\max(\ell, m, n) = \text{med}(\ell, m, n)$, we only use that $\sin(\cdot)$ and $\cos(\cdot)$ are bounded, which implies that

$$\begin{aligned} & \left| \sum_{\substack{\ell, m, n \in F_{\kappa, \kappa_0} : \\ \max(\ell, m, n) = \text{med}(\ell, m, n)}} \ell^{-1/2} m^{-1/2} n^{1/2} \right. \\ & \quad \left. \int_{x-t}^{x+t} dy \chi(y/\epsilon)^3 \sin((\ell-1)y) \sin((m-1)y) \cos(ny) \right| \\ & \lesssim_{\epsilon} \sum_{\substack{\ell, m, n \in F_{\kappa, \kappa_0} : \\ \max(\ell, m, n) = \text{med}(\ell, m, n)}} \ell^{-1/2} m^{-1/2} n^{1/2} \lesssim \sum_{\ell, m \in F_{\kappa, \kappa_0}} \ell^{-1/2} \lesssim 2^{-5\kappa_0} (\kappa - \kappa_0). \end{aligned}$$

This completes the proof of (A.6). \square

References

- [ADC21] Aizenman, M., Duminil-Copin, H.: Marginal triviality of the scaling limits of critical 4D Ising and ϕ_4^4 models. *Ann. Math. (2)* **194**(1), 163–235 (2021)
- [AD99] Andersson, L., Driver, B.K.: Finite-dimensional approximations to Wiener measure and path integral formulas on manifolds. *J. Funct. Anal.* **165**(2), 430–498 (1999)
- [BCD11] Bahouri, H., Chemin, J.-Y., Danchin, R.: *Fourier Analysis and Nonlinear Partial Differential Equations*. Grundlehren der Mathematischen Wissenschaften, vol. 343. Springer, Heidelberg (2011)
- [BOP15] Bényi, A., Oh, T., Pocovnicu, O.: On the probabilistic Cauchy theory of the cubic nonlinear Schrödinger equation on \mathbb{R}^d , $d \geq 3$. *Trans. Am. Math. Soc. Ser. B* **2**, 1–50 (2015)
- [BOP19] Bényi, A., Oh, T., Pocovnicu, O.: On the probabilistic Cauchy theory for nonlinear dispersive PDEs. In: *Landscapes of Time-Frequency Analysis. Applied and Numerical Harmonic Analysis*, pp. 1–32. Birkhäuser/Springer, Cham (2019)
- [Bou93] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.* **3**(2), 107–156 (1993)
- [Bou94] Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. *Commun. Math. Phys.* **166**(1), 1–26 (1994)

- [Bou96] Bourgain, J.: Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. *Commun. Math. Phys.* **176**(2), 421–445 (1996)
- [Bri18] Bringmann, B.: Almost sure scattering for the energy critical nonlinear wave equation. [arXiv:1812.10187](#). To appear in *Am. J. Math.* (2018)
- [Bri20] Bringmann, B.: Invariant Gibbs measures for the three-dimensional wave equation with a Hartree nonlinearity II: dynamics. [arXiv:2009.04616](#) (2020)
- [Bri21] Bringmann, B.: Almost sure local well-posedness for a derivative nonlinear wave equation. *Int. Math. Res. Not. IMRN* **2021**(11), 8657–8697 (2021)
- [BDNY22] Bringmann, B., Deng, Y., Nahmod, A. R., Yue, H.: Invariant Gibbs measures for the three dimensional cubic nonlinear wave equation. [arXiv:2205.03893](#) (2022)
- [BGHZ21] Bruned, Y., Gabriel, F., Hairer, M., Zambotti, L.: Geometric stochastic heat equations. *J. Am. Math. Soc.* **35**(1), 1–80 (2021)
- [BJ22] Brzeźniak, Z., Jendrej, J.: Statistical mechanics of the wave maps equation in dimension 1+1. [arXiv:2206.13605](#) (2022)
- [BR20] Brzeźniak, Z., Rana, N.: Low regularity solutions to the stochastic geometric wave equation driven by a fractional Brownian sheet. [arXiv:2006.07740](#) (2020)
- [BT08a] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. I. Local theory. *Invent. Math.* **173**(3), 449–475 (2008)
- [BT08b] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. II. A global existence result. *Invent. Math.* **173**(3), 477–496 (2008)
- [CM18] Cannizzaro, G., Matetski, K.: Space-time discrete KPZ equation. *Commun. Math. Phys.* **358**(2), 521–588 (2018)
- [CCHS20] Chandra, A., Chevyrev, I., Hairer, M., Shen, H.: Langevin dynamic for the 2D Yang-Mills measure. [arXiv:2006.04987](#) (2020)
- [CCHS22] Chandra, A., Chevyrev, I., Hairer, M., Shen, H.: Stochastic quantisation of Yang-Mills-Higgs in 3D. [arXiv:2201.03487](#) (2022)
- [CG19] Collot, C., Germain, P.: On the derivation of the homogeneous kinetic wave equation. [arXiv:1912.10368](#) (2019)
- [CLS21] Compagn, E., Lucà, R., Staffilani, G.: Pointwise convergence of the Schrödinger flow. *Int. Math. Res. Not. IMRN* **2021**(1), 599–650 (2021)
- [DPD02] Da Prato, G., Debussche, A.: Two-dimensional Navier-Stokes equations driven by a space-time white noise. *J. Funct. Anal.* **196**(1), 180–210 (2002)
- [DNY19] Deng, Y., Nahmod, A., Yue, H.: Invariant Gibbs measures and global strong solutions for nonlinear Schrödinger equations in dimension two. [arXiv:1910.08492](#) (2019)
- [DNY20] Deng, Y., Nahmod, A., Yue, H.: Random tensors, propagation of randomness, and nonlinear dispersive equations. [arXiv:2006.09285](#) (2020)
- [DH19] Deng, Y., Hani, Z.: On the derivation of the wave kinetic equation for NLS. [arXiv:1912.09518](#) (2019)
- [DH21] Deng, Y., Hani, Z.: Full derivation of the wave kinetic equation. [arXiv:2104.11204](#) (2021)
- [DNY21] Deng, Y., Nahmod, A.R., Yue, H.: Invariant Gibbs measure and global strong solutions for the Hartree NLS equation in dimension three. *J. Math. Phys.* **62**(3), 39 (2021)
- [DTV15] Deng, Y., Tzvetkov, N., Visciglia, N.: Invariant measures and long time behaviour for the Benjamin-Ono equation III. *Commun. Math. Phys.* **339**(3), 815–857 (2015)
- [DLM20] Dodson, B., Lührmann, J., Mendelson, D.: Almost sure scattering for the 4D energy-critical defocusing nonlinear wave equation with radial data. *Am. J. Math.* **142**(2), 475–504 (2020)
- [EH19] Erhard, D., Hairer, M.: Discretisation of regularity structures. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**(4), 2209–2248 (2019)
- [Eva13] Evans, L.C.: *An Introduction to Stochastic Differential Equations*. American Mathematical Society, Providence, RI (2013)
- [For20] Forlano, J.: Almost sure global well posedness for the BBM equation with infinite L^2 initial data. *Discret. Contin. Dyn. Syst.* **40**(1), 267–318 (2020)
- [FO20] Forlano, J., Okamoto, M.: A remark on norm inflation for nonlinear wave equations. *Dyn. Partial Differ. Equ.* **17**(4), 361–381 (2020)
- [FH17] Funaki, T., Hoshino, M.: A coupled KPZ equation, its two types of approximations and existence of global solutions. *J. Funct. Anal.* **273**(3), 1165–1204 (2017)
- [GJ87] Glimm, J., Jaffe, A.: *Quantum Physics*, second edition Springer-Verlag, New York (1987). (**A functional integral point of view**)
- [GIP15] Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, e6 (2015)
- [GKO18] Gubinelli, M., Koch, H., Oh, T.: Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. [arXiv:1811.07808](#). To appear in *J. Eur. Math. Soc.* (2018)

- [GH21] Gubinelli, M., Hofmanová, M.: A PDE construction of the Euclidean ϕ_3^4 quantum field theory. *Commun. Math. Phys.* **384**(1), 1–75 (2021)
- [Hai14] Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014)
- [HM18] Hairer, M., Matetski, K.: Discretisations of rough stochastic PDEs. *Ann. Probab.* **46**(3), 1651–1709 (2018)
- [Hai16] Hairer, M.: The motion of a random string. [arXiv:1605.02192](https://arxiv.org/abs/1605.02192) (2016)
- [HM12] Hairer, M., Maas, J.: A spatial version of the Itô–Stratonovich correction. *Ann. Probab.* **40**(4), 1675–1714 (2012)
- [Hsu02] Hsu, E.P.: *Stochastic Analysis on Manifolds*. Graduate Studies in Mathematics, vol. 38. American Mathematical Society, Providence, RI (2002)
- [IW89] Ikeda, N., Watanabe, S.: *Stochastic differential equations and diffusion processes*, volume 24 of North-Holland Mathematical Library. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, second edition (1989)
- [KS91] Karatzas, I., Shreve, S.E.: *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics, vol. 113, 2nd edn. Springer-Verlag, New York (1991)
- [KT98] Keel, M., Tao, T.: Local and global well-posedness of wave maps on \mathbb{R}^{1+1} for rough data. *Int. Math. Res. Not.* **1998**(21), 1117–1156 (1998)
- [KM19] Kenig, C., Mendelson, D.: The focusing energy-critical nonlinear wave equation with random initial data. [arXiv:1903.07246](https://arxiv.org/abs/1903.07246). To appear *Int. Math. Res. Not.* (2019)
- [KMV20] Killip, R., Murphy, J., Visan, M.: Invariance of white noise for KdV on the line. *Invent. Math.* **222**(1), 203–282 (2020)
- [Kis19] Kishimoto, N.: A remark on norm inflation for nonlinear Schrödinger equations. *Commun. Pure Appl. Anal.* **18**(3), 1375–1402 (2019)
- [KM93] Klainerman, S., Machedon, M.: Space-time estimates for null forms and the local existence theorem. *Commun. Pure Appl. Math.* **46**(9), 1221–1268 (1993)
- [KM95] Klainerman, S., Machedon, M.: Smoothing estimates for null forms and applications. *Duke Math. J.* **81**(1), 99–133 (1996), 1995
- [KM97] Klainerman, S., Machedon, M.: On the regularity properties of a model problem related to wave maps. *Duke Math. J.* **87**(3), 553–589 (1997)
- [KR01] Klainerman, S., Rodnianski, I.: On the global regularity of wave maps in the critical Sobolev norm. *Int. Math. Res. Not.* **2001**(13), 655–677 (2001)
- [KS97] Klainerman, S., Selberg, S.: Remark on the optimal regularity for equations of wave maps type. *Commun. Partial Differ. Equ.* **22**(5–6), 901–918 (1997)
- [KS02] Klainerman, S., Selberg, S.: Bilinear estimates and applications to nonlinear wave equations. *Commun. Contemp. Math.* **4**(2), 223–295 (2002)
- [Kri03] Krieger, J.: Global regularity of wave maps from \mathbb{R}^{3+1} to surfaces. *Commun. Math. Phys.* **238**(1–2), 333–366 (2003)
- [Kri04] Krieger, J.: Global regularity of wave maps from \mathbb{R}^{2+1} to \mathbb{H}^2 . Small energy. *Commun. Math. Phys.* **250**(3), 507–580 (2004)
- [KS12] Krieger, J., Schlag, W.: *Concentration Compactness for Critical Wave Maps*. EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich (2012)
- [KST08] Krieger, J., Schlag, W., Tataru, D.: Renormalization and blow up for charge one equivariant critical wave maps. *Invent. Math.* **171**(3), 543–615 (2008)
- [KLS20] Krieger, J., Lührmann, J., Staffilani, G.: Probabilistic small data global well-posedness of the energy-critical Maxwell–Klein–Gordon equation. [arXiv:2010.09528](https://arxiv.org/abs/2010.09528) (2020)
- [Kup16] Kupiainen, A.: Renormalization group and stochastic PDEs. *Ann. Henri Poincaré* **17**(3), 497–535 (2016)
- [LRS88] Lebowitz, J., Rose, H., Speer, E.: Statistical mechanics of the nonlinear Schrödinger equation. *J. Stat. Phys.* **50**(3–4), 657–687 (1988)
- [LM14] Lührmann, J., Mendelson, D.: Random data Cauchy theory for nonlinear wave equations of power-type on \mathbb{R}^3 . *Commun. Partial Differ. Equ.* **39**(12), 2262–2283 (2014)
- [MNT10] Machihara, S., Nakanishi, K., Tsugawa, K.: Well-posedness for nonlinear Dirac equations in one dimension. *Kyoto J. Math.* **50**(2), 403–451 (2010)
- [NSU03] Nahmod, A., Stefanov, A., Uhlenbeck, K.: On the well-posedness of the wave map problem in high dimensions. *Commun. Anal. Geom.* **11**(1), 49–83 (2003)
- [NORBS12] Nahmod, A.R., Oh, T., Rey-Bellet, L., Staffilani, G.: Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS. *J. Eur. Math. Soc. (JEMS)* **14**(4), 1275–1330 (2012)
- [NS19] Nahmod, A.R., Staffilani, G.: Randomness and nonlinear evolution equations. *Acta Math. Sin. (Engl. Ser.)* **35**(6), 903–932 (2019)

- [Nua06] Nualart, D.: The Malliavin calculus and related topics. Probability and its Applications (New York), 2nd edn. Springer-Verlag, Berlin (2006)
- [Oh17] Oh, T.: A remark on norm inflation with general initial data for the cubic nonlinear Schrödinger equations in negative Sobolev spaces. *Funkcial. Ekvac.* **60**(2), 259–277 (2017)
- [OOT20] Oh, T., Okamoto, M., Tolomeo, L.: Focusing Φ_3^4 -model with a Hartree-type nonlinearity. [arXiv:2009.03251](https://arxiv.org/abs/2009.03251) (2020)
- [OOT21] Oh, T., Okamoto, M., Tolomeo, L.: Stochastic quantization of the Φ_3^3 -model. [arXiv:2108.06777](https://arxiv.org/abs/2108.06777) (2021)
- [OST21] Oh, T., Sosoe, P., Tolomeo, L.: Optimal integrability threshold for Gibbs measures associated with focusing NLS on the torus. [arXiv:1709.02045](https://arxiv.org/abs/1709.02045) (2021)
- [OW19] Otto, F., Weber, H.: Quasilinear SPDEs via rough paths. *Arch. Ration. Mech. Anal.* **232**(2), 873–950 (2019)
- [Poh76] Pohlmeier, K.: Integrable Hamiltonian systems and interactions through quadratic constraints. *Commun. Math. Phys.* **46**(3), 207–221 (1976)
- [RR12] Raphaël, P., Rodnianski, I.: Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems. *Publ. Math. Inst. Hautes Études Sci.* **115**, 1–122 (2012)
- [RS10] Rodnianski, I., Sterbenz, J.: On the formation of singularities in the critical $O(3)$ σ -model. *Ann. Math. (2)* **172**(1), 187–242 (2010)
- [SS02] Shatah, J., Struwe, M.: The Cauchy problem for wave maps. *Int. Math. Res. Not.* **2002**(11), 555–571 (2002)
- [She21] Shen, H.: Stochastic quantization of an Abelian Gauge theory. *Commun. Math. Phys.* **384**(3), 1445–1512 (2021)
- [Sim96] Simon, L.: Theorems on Regularity and Singularity of Energy Minimizing Maps. *Lectures in Mathematics, ETH Zürich*. Birkhäuser Verlag, Basel (1996)
- [ST10a] Sterbenz, J., Tataru, D.: Energy dispersed large data wave maps in $2 + 1$ dimensions. *Commun. Math. Phys.* **298**(1), 139–230 (2010)
- [ST10b] Sterbenz, J., Tataru, D.: Regularity of wave-maps in dimension $2 + 1$. *Commun. Math. Phys.* **298**(1), 231–264 (2010)
- [Str03] Struwe, M.: Equivariant wave maps in two space dimensions. *Commun. Pure Appl. Math.* **56**(7), 815–823 (2003)
- [ST20] Sun, C., Tzvetkov, N.: Concerning the pathological set in the context of probabilistic well-posedness. *C. R. Math. Acad. Sci. Paris* **358**(9–10), 989–999 (2020)
- [ST21] Sun, C., Tzvetkov, N.: Refined probabilistic global well-posedness for the weakly dispersive NLS. *Nonlinear Anal.* **213**, 112530 (2021)
- [Tao01a] Tao, T.: Global regularity of wave maps. I. Small critical Sobolev norm in high dimension. *Int. Math. Res. Not.* **2001**(6), 299–328 (2001)
- [Tao01b] Tao, T.: Global regularity of wave maps. II. Small energy in two dimensions. *Commun. Math. Phys.* **224**(2), 443–544 (2001)
- [Tao08] Tao, T.: Global regularity of wave maps. III–VII. *arXiv preprints*, (2008+)
- [Tao00] Tao, T.: Ill-posedness for one-dimensional wave maps at the critical regularity. *Am. J. Math.* **122**(3), 451–463 (2000)
- [Tao06] Tao, T.: Nonlinear dispersive equations, volume 106 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI. Local and global analysis (2006)
- [Tat01] Tataru, D.: On global existence and scattering for the wave maps equation. *Am. J. Math.* **123**(1), 37–77 (2001)
- [Tat05] Tataru, D.: Rough solutions for the wave maps equation. *Am. J. Math.* **127**(2), 293–377 (2005)
- [Tat98] Tataru, D.: Local and global results for wave maps. I. *Commun. Partial Differ. Equ.* **23**(9–10), 1781–1793 (1998)
- [TU04] Terng, C.-L., Uhlenbeck, K.: $1 + 1$ Wave maps into symmetric spaces. *Commun. Anal. Geom.* **12**(1–2), 345–388 (2004)
- [Tzv15] Tzvetkov, N.: Quasiinvariant Gaussian measures for one-dimensional Hamiltonian partial differential equations (2015)