

# Almost sure boundedness of iterates for derivative nonlinear wave equations

SAGUN CHANILLO, MAGDALENA CZUBAK, DANA MENDELSON,  
ANDREA NAHMOD, AND GIGLIOLA STAFFILANI

*In celebration of Karen Uhlenbeck’s 75th birthday*

We study nonlinear wave equations on  $\mathbb{R}^{2+1}$  with quadratic derivative nonlinearities, which include in particular nonlinearities exhibiting a null form structure, with random initial data in  $H_x^1 \times L_x^2$ . In contrast to the counterexamples of Zhou [73] and Foschi-Klainerman [23], we obtain a uniform time interval  $I$  on which the Picard iterates of all orders are almost surely bounded in  $C_t(I; \dot{H}_x^1)$ .

<b>1 Introduction</b>	<b>942</b>
<b>2 Set-up and statement of main results</b>	<b>947</b>
<b>3 Preliminaries</b>	<b>949</b>
<b>4 The iteration scheme</b>	<b>953</b>
<b>5 Proof of main probabilistic bounds</b>	<b>960</b>
<b>6 Proof of Theorem 2.1</b>	<b>967</b>
<b>References</b>	<b>968</b>

S.C. is funded in part by NSF DMS-1201474.

M.C. is funded in part by the Simons Foundation #246255.

D.M. was funded in part by NSF DMS-1128155 during the completion of this work.

A.N. is funded in part by NSF DMS-1201443 and DMS-1463714.

G.S. is funded in part by NSF DMS-1362509, DMS-1462401, John Simon Guggenheim Foundation, and the Simons Foundation.

## 1. Introduction

We consider systems of derivative nonlinear wave equations given by

$$(1.1) \quad \square u^I = \sum_{J,K} \partial u^J \partial u^K,$$

where we use  $\partial$  to denote a first order derivative, i.e.  $\partial \in \{\partial_t, \partial_{x_1}, \partial_{x_2}\}$ . In particular, this class of equations includes the so-called null-form derivative nonlinear wave equations

$$(1.2) \quad \square u^I = \sum_{J,K} Q(u^J, u^K),$$

where  $\square = -\partial_t^2 + \Delta$ ,  $(u^I) : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}^m$ ,  $m \in 2, 3, \dots$  and  $Q$  is a bilinear form given by

$$\begin{aligned} Q_0(f, g) &= \partial_t f \partial_t g - \nabla f \cdot \nabla g, \\ Q_{0j}(f, g) &= \partial_t f \partial_j g - \partial_j f \partial_t g, \\ Q_{ij}(f, g) &= \partial_i f \partial_j g - \partial_j f \partial_i g, \end{aligned}$$

where  $\partial_j$  stands for the spatial derivatives and  $\nabla$  for the spatial gradient.

We now give some history of the wave equations with null forms as well as of the more general derivative nonlinear wave equations. The null forms were introduced by Klainerman [29, 30] (see also Christodoulou [13]) and nonlinearities exhibiting quadratic derivatives with such null form structure appear in many physical models and geometric wave equations, such as wave maps, the Maxwell-Klein-Gordon system, the Yang-Mills equations, the space-time Monopole equation, and Ward wave maps.

Wave maps are one of the simplest geometric wave equations and can be viewed as the Minkowski’s analogue of a harmonic map. We refer the reader, for example to [14, 37, 39, 48, 55, 56, 59, 60, 62] for some of the pioneering works. The equations are particularly interesting in the two dimensional case which is energy critical.

The Maxwell-Klein-Gordon system can be viewed as an abelian analogue of Yang-Mills. Fundamental contributions to the mathematical analysis of the Yang-Mills gauge theory were made by Uhlenbeck [65, 66], see also the book of Freed and Uhlenbeck [24] for further discussion. We also refer the reader to [17, 32, 38, 40–42, 47, 49, 51] and references therein for more on the Maxwell-Klein-Gordon and Yang-Mills equations.

The Ward wave map and the space-time Monopole equation were introduced by Ward in [67, 68]. The first one is an integrable system in  $2 + 1$  dimensions while the second one is a space-time analogue of Bogomolny equations. The Monopole equation can be derived from the anti-self dual Yang-Mills by a dimensional reduction, and the Ward wave map from the Monopole equation by choosing a particular gauge. For results on the Cauchy problem for Ward wave maps and soliton construction see [15, 18, 26, 69]. A broad survey on the space-time Monopole equation is given by Dai, Terng and Uhlenbeck [19], where, in particular, they show global existence and uniqueness up to a gauge transformation for small initial data in  $W^{2,1}$  via a scattering transform. For results on the Cauchy problem for the Monopole equation, see [8, 16, 63].

Our goal in this note is to study (1.1) from a probabilistic point of view. Our motivation for this is two-fold. First, as we will see when we detail the deterministic local well-posedness theory below, in two dimensions there exists a gap in Sobolev space between the scaling invariant space and the optimal local well-posedness for general quadratic derivative nonlinearities (1.1), including null-form nonlinear wave equations with  $Q_{ij}$  and  $Q_{0j}$  nonlinearities. In this work we make progress in bridging this gap. A second reason we have for studying (1.1) is that it serves as a model problem for (1.2) and the geometric flows described above, and hence this note fits into a broader program of adapting probabilistic techniques to geometric equations. In many of these cases, however, one would need to define a new randomization procedure which respects the geometry of the target manifold, is compatible with gauge transformations and pull-backs, and which still yields (almost surely) better properties for the solutions. This is largely uncharted territory and represents an exciting future direction of research which will require many new ideas.

We review the well-posedness results for (1.1) and (1.2). Schematically, we compare these to the following derivative nonlinear wave equation:

$$(1.3) \quad \square u = (\partial u)^2,$$

where  $u : \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ , and  $\partial$  is any of the derivatives  $\partial_\alpha$ ,  $\alpha = 0, \dots, d$ . Solutions to (1.3) are invariant under the scaling transformation

$$u \mapsto u_\lambda(t, x) = u(\lambda t, \lambda x)$$

and the scaling critical regularity,  $s_c = \frac{d}{2}$ , is by definition the regularity such that the corresponding homogeneous Sobolev space  $\dot{H}^{s_c}(\mathbb{R}^d)$  is invariant

under the above scaling. In particular, in dimension  $d = 2$ , one has  $s_c = 1$ , and hence (1.3) is energy critical. Energy methods yield local well-posedness for  $s > \frac{d}{2} + 1$ , and using Strichartz estimates (see [53]) one may improve this to  $s > \max(\frac{d}{2}, \frac{d+5}{4})$ , which is sharp for (1.3) with data in the Sobolev class in light of the counterexamples of Lindblad [44]. Nonetheless, this leaves a gap in dimensions  $d = 2, 3, 4$  between the optimal regularity in Sobolev spaces and the scaling prediction. In dimension  $d = 3$ , this gap was closed for data in the Fourier-Lebesgue class [28]. See also [20, 27, 61] for improved estimates in 2D.

By replacing the general product  $(\partial u)^2$  by one of the null forms, Klainerman and Machedon demonstrated in [31] that one may lower the regularity required for local well-posedness from  $s > 2$  to  $s = 2$  for either  $Q_0$ ,  $Q_{0j}$  or  $Q_{ij}$  in dimension  $d = 3$ . Subsequently, it was shown in [33, 35, 36] that in the particular case of (1.2) with a  $Q_0$  nonlinearity local well-posedness holds in  $H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$  with  $s > \frac{d}{2}$  and  $d \geq 2$ , which is almost optimal. Similarly, for the  $Q_{0j}$  and  $Q_{ij}$  null form nonlinearities, almost optimal local well-posedness can also be achieved in Sobolev spaces with index  $s > \frac{d}{2}$  but in dimensions  $d \geq 3$  [34].

In the energy critical setting  $d = 2$ , the situation is more delicate for the  $Q_{0j}$  and  $Q_{ij}$  null form nonlinearities. The best known result in Sobolev spaces was achieved by Zhou [73], who proved local well-posedness for  $s > \frac{5}{4}$ , which we note is still  $\frac{1}{4}$  away from the scaling critical regularity. Examination of the first iterate shows that in Sobolev spaces, this is optimal using iteration methods, see [23, 36, 73] for further discussion. This gap in dimension  $d = 2$  was closed for data in the Fourier-Lebesgue class in [26].

As mentioned above, our aim is to close this gap between the local well-posedness theory and the scaling critical regularity in Sobolev spaces with the aid of probabilistic methods. The use of probabilistic techniques to study the well-posedness of nonlinear dispersive equations was initiated by Bourgain motivated by the question of the invariance of associated Gibbs measures<sup>1</sup>. In [5, 6] Bourgain studied the periodic nonlinear Schrödinger equation in one and two space dimensions, in [7] the Gross-Pitaevskii equation and in [4] the Zakharov system. In particular for the periodic cubic nonlinear Schrödinger equation in two dimensions, Bourgain constructed, for the first time, sets of supercritical initial data of close to full measure which give rise to local in time solutions, and then proved almost sure global well-posedness via the invariance of the Gibbs measure.

<sup>1</sup>after the works by Lebowitz, Rose and Speer [43], by Glimm and Jaffe [25] and by Zhidkov [71, 72]

Following the seminal contributions of Bourgain, extensive progress has been made in recent years in the study of random data well-posedness of nonlinear dispersive, wave and fluid equations, both in compact and non-compact settings. In the context of nonlinear wave equations, and after work by Zhidkov [71], Burq and Tzvetkov considered in [10, 11] the cubic nonlinear wave equation on a three-dimensional compact manifold. As in the two dimensional result of Bourgain, Burq and Tzvetkov also constructed sets of supercritical initial data of close to full measure which give rise to local solutions, and subsequently proved the almost sure existence of global solutions for the radial problem via the invariance of the associated Gibbs measure. For other probabilistic results on nonlinear wave equations see for instance [9, 12, 21, 45, 46, 50, 52, 57, 58] and references therein.

As the randomization of the initial data does not regularize in Sobolev spaces, the free evolution of the random initial data is almost surely no more regular than the function which was randomized. Consequently, the typical scheme employed for the random data local well-posedness theory in previous works on nonlinear wave and dispersive equations with power-type nonlinearities has involved re-centering the flow and proving a fixed point argument for the nonlinear component of the solution. Carrying out this method requires one to gain regularity for the nonlinear component of the solution, and it is not clear if this can be done in the current setting of a quadratic derivative nonlinearity. We also note that it is sometimes possible to further re-center the initial data in cases where regularity is only gained for certain components of the nonlinearity, see for instance the recent work of [3].

### Overview of the main results

We state the main theorem precisely in Section 2. Here we just give the main idea. We study a scalar version of equation (1.1) given by

$$(1.4) \quad \square u = (\partial u)^2.$$

We consider a pair of functions  $(\phi_0, \phi_1) \in H^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ , and we randomize them according to a unit-scale projection in frequency space. See Section 2 for the precise definition of the randomization. We then consider the Picard iterates for (1.4), and show that the  $n$ -th iterate,  $u^{(n)}(x, t)$ , is almost surely bounded, that is, there exists  $T$  sufficiently small such that

946 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

with probability one, we have

$$u^{(n)}(x, t) \in C_t([0, T]; \dot{H}^1(\mathbb{R}^d)),$$

for every  $n \geq 0$ ; see Theorem 2.1 for the precise statement. The analogous result can be also obtained for vector-valued equations. See Remark 2.2

This result should be held in contrast to the deterministic results of Zhou which indicate that for initial data  $\varphi \in H^s(\mathbb{R}^2)$ , there is no estimate for the first iterate for  $s \leq \frac{5}{4}$ . More precisely, in [73, Proposition 5] Zhou proves that given functions  $f$  and  $g$  which lie in  $H^{s+1}(\mathbb{R}^2)$  for  $0 \leq s \leq \frac{1}{4}$ , if we consider

$$F = \cos(t|\nabla|)f, \quad G = \cos(t|\nabla|)g$$

then the solution  $\psi$  to

$$\begin{aligned} \square\psi &= Q_{12}(F, G), \\ (\psi, \psi_t)|_{t=0} &= (0, 0), \end{aligned}$$

fails to be in  $H^{s+1}(\mathbb{R}^2)$ , and in particular, the estimate

$$\|\psi(t, \cdot)\|_{H^{s+1}} + \|\partial_t \psi(t, \cdot)\|_{H^s} \leq C(t)\|f\|_{H^{s+1}}\|g\|_{H^{s+1}}$$

fails.

### Notation and conventions

We denote by  $C > 0$  an absolute constant that depends only on fixed parameters and whose value may change from line to line. We write  $X \lesssim Y$  if  $X \leq CY$  for some  $C > 0$ , and analogously for  $X \gtrsim Y$ . Finally, we will use  $\mathcal{F}f$  or  $\widehat{f}$  to denote the Fourier transform of a function  $f$ .

### Organization of paper

In Section 2, we introduce our framework and give a precise statement of the main result, Theorem 2.1. In Section 3, we collect some probabilistic facts which we will need in our analysis. In Section 4, we set up the iteration scheme. In Section 5, we establish the main probabilistic bounds for the iterates, and finally, in Section 6, we prove the main theorem, Theorem 2.1.

## 2. Set-up and statement of main results

We introduce the randomization procedure by which we construct the initial data. We will present the randomization for real-valued initial data which readily generalizes to targets  $\mathbb{R}^m$  by randomizing each coordinate according to the procedure described below.

To define the randomized initial data on Euclidean space, we let  $\psi \in C_c^\infty(\mathbb{R}^2)$  be a non-negative function with  $\text{supp}(\psi) \subset B(0, 1)$  and such that

$$\sum_{k \in \mathbb{Z}^2} \psi(\xi - k) = 1 \text{ for all } \xi \in \mathbb{R}^2.$$

For every  $k \in \mathbb{Z}^2$ , we define the function  $P_k f : \mathbb{R}^2 \rightarrow \mathbb{C}$  for any  $f \in H^s(\mathbb{R}^2)$  by

$$(P_k f)(x) = \mathcal{F}^{-1} \left( \psi(\xi - k) \hat{f}(\xi) \right) (x) \text{ for } x \in \mathbb{R}^2.$$

We set  $\phi = (\phi_0, \phi_1) \in H^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$  and we let  $\{(\varepsilon_k, \nu_k)\}_{k \in \mathbb{Z}^2}$  be a sequence of identically distributed independent Rademacher random variables on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . We recall that by definition these are centered random variables which take values  $\pm 1$  with equal probability. Define

$$(2.1) \quad \phi^\omega = (\phi_0^\omega(x), \phi_1^\omega(x)) := \left( \sum_{k \in \mathbb{Z}^2} \varepsilon_k(\omega) P_k \phi_0(x), \sum_{k \in \mathbb{Z}^2} \nu_k(\omega) P_k \phi_1(x) \right),$$

where this quantity is understood as a Cauchy limit in  $L^2(\Omega; H^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2))$ . Similar randomizations have previously been used in [70], [45], [2]. Crucially, such a randomization does not regularize at the level of Sobolev spaces, see [10].

### 2.1. Statement of main results

We denote the free wave evolution of the initial data  $\phi^\omega$  by

$$(2.2) \quad W(t)\phi^\omega = \cos(t|\nabla|)\phi_0^\omega + \frac{\sin(t|\nabla|)}{|\nabla|}\phi_1^\omega.$$

Since the randomization does not regularize in Sobolev spaces, the free evolution of the random initial data (2.2) remains at regularity  $H^1(\mathbb{R}^2) \times$

948 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

$L^2(\mathbb{R}^2)$ . Typically, when studying such problem, one writes solutions as

$$(u(t), \partial_t u(t)) = (W(t)\phi^\omega, \partial_t W(t)\phi^\omega) + (w(t), \partial_t w(t)),$$

and proves that the nonlinear component  $(w(t), \partial_t w(t))$  is almost surely smoother than the random initial data (this dates back to [6]). In studying (1.4), due to the derivative structure of the nonlinearity, it is not clear that the solution lies in a smoother space. We instead turn to the Picard iterates. We define

$$(2.3) \quad u^{(0)}(x, t) = W(t)(\phi_0^\omega, \phi_1^\omega)$$

and, setting the notation

$$A_0(\cdot, \cdot) = \int_0^t \frac{\partial \sin((t-t')|\nabla|)}{|\nabla|}(\cdot)(\cdot) dt',$$

we define the  $n$ -th iterate inductively as

$$(2.4) \quad \begin{aligned} \partial u^{(n)} &= \partial u^{(0)} + A_0(\partial u^{(n-1)}, \partial u^{(n-1)}), \\ \partial u^{(0)} &= \sum_k \varepsilon_k(\omega) P_k \partial W(t) \phi =: \sum_k \varepsilon_k(\omega) F_k. \end{aligned}$$

Our goal is to prove that almost surely, the iterates for (1.4) are bounded. We are now prepared to state our main theorem.

**Theorem 2.1.** *Let  $(\Omega, \mathcal{A}, \mathbb{P})$  be a probability space. Let  $(\phi_0, \phi_1) \in H^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$  and let  $\phi^\omega$  be as given in (2.1). Let  $u^{(0)}$  be the free evolution of  $\phi^\omega$  defined in (2.3) and let the  $n$ -th Picard iterate be as defined in (2.4). Then for any  $T > 0$  sufficiently small, there exists  $\Sigma_T \subseteq \Omega$ , with  $\mathbb{P}(\Sigma_T) = 1$ , such that for every  $\omega \in \Sigma_T$  and  $n \geq 0$ ,*

$$(u^{(n)}, \partial_t u^{(n)}) \in C_t([0, T]; \dot{H}^1(\mathbb{R}^2) \times L^2(\mathbb{R}^2)).$$

**Remark 2.2.** Although we prove this for scalar valued equations, the results generalize to vector-valued equations by considering the terms component-wise. For general initial data with an  $m$ -dimensional target, we consider a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  and we let  $\{(\varepsilon_k, \nu_k)\}_{k \in \mathbb{Z}^2}$  be a sequence of  $m$ -dimensional Rademacher random variables. We then randomize the initial data with multi-dimensional target by randomizing each coordinate separately. In particular, Theorem 2.1 yields the same results for systems of derivative nonlinear wave equations, as in (1.1).



Our proof relies on writing the  $n$ -th iterate in terms of not only the previous iterate, but on *un-packing* it all the way to the free solution. Because the nonlinearity is quadratic, this results in  $2^n$  many unit-sized projections of the free solutions interacting together at the  $n$ -th iterate level. By expanding the iterates completely, the randomization will then enable us to use the Bernstein inequality of Lemma 3.1 below to take advantage of the support properties of each term. We will also show that the quadratic structure naturally gives rise to a connection with binary trees.

**Remark 2.3.** An identical proof will yield Theorem 2.1 for dimensions three and four. The only modification would be the factor appearing in Proposition 4.9 arising from Bernstein’s inequality, see Remark 4.10 for details and for why dimensions two, three and four are favorable.

In dimension three, the critical Sobolev exponent is  $\frac{3}{2}$ , and in dimension four, the critical exponent is 2. Because the theorem applies to data in  $H^1 \times L^2$ , and the iterates  $(u^{(n)}, \partial_t u^{(n)})$  belong to  $C_t([0, T]; \dot{H}^1 \times L^2)$ , this means the result is in the supercritical regime in dimensions three and four. Finally, we note that as we only use the energy estimate on the inhomogeneous term, the Strichartz estimates are not needed for any of these dimensions.

### Acknowledgements

The second (M.C.) and fourth (A.N.) authors would like to warmly thank Karen Uhlenbeck for inspiring us, and for all the mathematics that she has taught us and continues to teach us.

The authors thank the MSRI, the IHES and the Radcliffe Institute for Advanced Study at Harvard University for the kind hospitality that allowed us to develop this project. D.M. gratefully acknowledges support from the Institute for Advanced Study at Princeton, and would also like to thank Carlos Kenig for helpful conversations.

### 3. Preliminaries

Here we record some facts which will be of use to us in our analysis. The first is a unit-scale Bernstein estimate for the projection operators.

**Lemma 3.1 (Bernstein inequality).** *Let  $1 \leq q \leq 2 \leq p \leq \infty$ . There exists a constant  $C_0$ , depending only on the dimension  $n$ , and  $p, q$  such that if  $f$  is a function with a support of  $\hat{f}$  contained in a measurable set  $E$ , then*

$$\|f\|_{L_x^p} \leq C_0 |E|^{\frac{1}{q} - \frac{1}{p}} \|f\|_{L_x^q}.$$

**Remark 3.2.** We apply this in two situations. When we have a single function  $P_k f$ , and when we have a product of  $j$  functions  $P_{k_i} f$  for  $i = 1, \dots, j$  and  $j \geq 2$ . In both of these cases, we will take  $p = 4$  and  $q = 2$ . This leads to

$$\|P_k f\|_{L_x^4} \leq C_0 \pi^{\frac{1}{4}} \|P_k f\|_{L_x^2},$$

and

$$\|P_{k_1} f \cdots P_{k_j} f\|_{L_x^4} \leq C_0 \pi^{\frac{1}{4}} j^{\frac{1}{2}} \|P_k f\|_{L_x^2},$$

since the supports are then contained in the ball of radius 1 and  $j$ , respectively.

In the sequel, it will be useful to introduce the notation

$$(3.1) \quad M_{01} := \frac{\partial \sin((t_0 - t_1)|\nabla|)}{|\nabla|},$$

where once again we mention that  $\partial$  can be either a spatial derivative or a time derivative. We will repeatedly make use of the following fact about the multiplier  $M_{01}$ .

**Lemma 3.3.** *The multiplier  $M_{01}$  is bounded on  $L_x^2$ , with*

$$\|M_{01} f\|_{L_x^2} \leq \|f\|_{L_x^2}.$$

### 3.1. Large deviation estimates

We begin by recalling a large deviation estimate, which goes back to the classical work of Kolmogorov, Paley and Zygmund.

**Lemma 3.4** ([10, Lemma 3.1]). *Let  $\{\varepsilon_k\}_{k=1}^\infty$  be a sequence of independent identically distributed (iid) Rademacher random variables on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Then there exists  $C > 0$  such that for every  $p \geq 2$  and every  $\{c_k\}_{k=1}^\infty \in \ell^2(\mathbb{N}; \mathbb{C})$ , we have*

$$(3.2) \quad \left\| \sum_{k=1}^{\infty} c_k \varepsilon_k(\omega) \right\|_{L_\omega^p(\Omega)} \leq C \sqrt{p} \left( \sum_{k=1}^{\infty} |c_k|^2 \right)^{\frac{1}{2}}.$$

As a consequence of Chebychev’s inequality, there exists  $\alpha > 0$  such that for every  $\lambda > 0$  and every sequence  $\{c_k\}_{k=1}^\infty \in \ell^2(\mathbb{N}; \mathbb{C})$  of complex numbers,

$$\mathbb{P} \left( \left\{ \omega : \left| \sum_{k=1}^\infty c_k \varepsilon_k(\omega) \right| > \lambda \right\} \right) \leq 2 \exp \left( -\alpha \frac{\lambda^2}{\sum_k |c_k|^2} \right).$$

We will also use the following lemma, which can be viewed as a generalization of estimate (3.2) which allows for nondeterministic coefficients. A version of this lemma appeared in [22], however we include here a proof of this fact for completeness.

**Lemma 3.5.** *Let  $\{\varepsilon_k\}_{k=1}^\infty$  be a sequence of iid Rademacher random variables on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ , and let  $b_k$  be a sequence of random variables which are independent of the  $\{\varepsilon_k\}_{k=1}^\infty$ . Then there exists  $C > 0$  such that for all  $p \geq 2$ ,*

$$(3.3) \quad \left\| \sum_{k=1}^\infty \varepsilon_k b_k \right\|_{L_\omega^p(\Omega)} \leq C \sqrt{p} \left\| \left( \sum_{k=1}^\infty |b_k|^2 \right)^{\frac{1}{2}} \right\|_{L_\omega^p(\Omega)}.$$

*Proof.* First we observe that by the Monotone Convergence Theorem and the Cauchy criterion for convergence of infinite series, it is enough to prove the estimate for a finite sum.

Next we will prove the desired inequality for even powers of  $p$ , i.e.,  $p = 2j \geq 2$ . Using the independence and that  $\mathbb{E}(\varepsilon_k) = 0$ , one can easily see that

$$(3.4) \quad \left\| \sum_{k=1}^N \varepsilon_k b_k \right\|_{L_\omega^p(\Omega)}^p = \sum_{2k_1 + \dots + 2k_N = 2j} \int \frac{(2j)!}{(2k_1)! \dots (2k_N)!} |b_1|^{2k_1} \dots |b_N|^{2k_N}.$$

Expanding the right hand side of (3.3), we similarly get for  $p = 2j$  that

$$(3.5) \quad \left\| \left( \sum_{k=1}^N |b_k|^2 \right)^{\frac{1}{2}} \right\|_{L_\omega^p(\Omega)}^p = \sum_{k_1 + \dots + k_N = j} \int \frac{(j)!}{k_1! \dots k_N!} |b_1|^{2k_1} \dots |b_N|^{2k_N}.$$

Thus, comparing the right hand side of (3.4) with (3.5) gives us an estimate with constant

$$(3.6) \quad \max_{j_1, \dots, j_N} \frac{(2j)!}{j!} \frac{k_1! \dots k_N!}{(2k_1)! \dots (2k_N)!}.$$

952 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

However

$$\frac{k_1! \cdots k_N!}{(2k_1)! \cdots (2k_N)!} \leq 1,$$

so an application of the Stirling’s formula yields then a constant of

$$\frac{(2j)!}{j!} = C^{2j} j^j.$$

Noting that the bound does not depend on  $N$ , the desired estimate follows in the  $p = 2j$  case.

To obtain (3.3) for arbitrary  $p \geq 2$ , note first that for  $p = 2j + 2$  one obtains a similar estimate with  $C(2j + 2)^{1/2}$  for the same constant  $C$  appearing above. By interpolation for mixed-norm spaces, see [1, Section 7, Theorem 2] one obtains the desired inequality (3.3) for  $2j < p < (2j + 2)$  with constant  $\sqrt{2}C\sqrt{p}$ , which yields the desired bound for all  $p \geq 2$ .  $\square$

We will also use the following variant of Lemma 4.5 in [64] to bound the probability of certain subsets of the probability space.

**Lemma 3.6.** *Let  $F$  be a real valued measurable function on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Suppose that there exists  $\alpha > 0$ ,  $N > 0$ ,  $k \in \mathbb{N}^*$  and  $C > 0$  such that for every  $p \geq p_0 \geq 1$  one has*

$$\|F\|_{L^p_\omega(\Omega)} \leq CN^{-\alpha} p^{\frac{k}{2}}.$$

*Then, there exists  $C_1$ , and  $c$  depending on  $C$  and  $p_0$  such that for  $\lambda > 0$*

$$\mathbb{P}(\omega \in \Omega : |F(\omega)| > \lambda) =: \mathbb{P}(E_\lambda) \leq C_1 e^{-cN^{\frac{2\alpha}{k}} \lambda^{\frac{2}{k}}}.$$

We now state the improved linear estimate for the zeroth iterates.

**Proposition 3.7.** *Let  $\phi^\omega$  be as defined in (2.1), and  $u^{(0)}(x, t)$  the zeroth iterate defined in (2.3). Let  $2 \leq q, r < \infty$ . Then*

$$\|\partial u^{(0)}\|_{L^q_t L^r_x(I \times \mathbb{R}^2)} < \infty$$

*almost surely.*

*Proof.* We have

$$\partial u^{(0)}(t) = \partial W(t)\phi^\omega = \partial \cos(t|\nabla|)\phi_0^\omega + \partial \frac{\sin(t|\nabla|)}{|\nabla|}\phi_1^\omega.$$

We only prove the estimate for  $\partial e^{\pm it|\nabla|}\phi_0^\omega$  since the other terms, including the term involving  $\phi_1^\omega$ , follow analogously. Let  $p \geq \max(q, r)$ . Then by Lemma 3.4 and Minkowski's inequality

$$\begin{aligned} \|\partial e^{\pm it|\nabla|}\phi_0^\omega\|_{L_t^q L_x^r(I \times \mathbb{R}^2)}\|_{L_\omega^p} &\leq \|\partial e^{\pm it|\nabla|}\phi_0^\omega\|_{L_\omega^p}\|_{L_t^q L_x^r(I \times \mathbb{R}^2)} \\ &\lesssim \sqrt{p} \left( \sum_k \|\partial e^{\pm it|\nabla|} P_k \phi_0\|_{L_t^q L_x^r(I \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

We then use Hölder's inequality and the unit-scale Bernstein estimate of Lemma 3.1 to obtain

$$\begin{aligned} \|\partial e^{\pm it|\nabla|}\phi_0^\omega\|_{L_t^q L_x^r(I \times \mathbb{R}^2)}\|_{L_\omega^p} &\lesssim \sqrt{p} |I|^{\frac{1}{q}} \left( \sum_k \|\partial e^{\pm it|\nabla|} P_k \phi_0\|_{L_t^\infty L_x^2(I \times \mathbb{R}^2)}^2 \right)^{\frac{1}{2}} \\ &\lesssim \sqrt{p} |I|^{\frac{1}{q}} \left( \sum_k \|\nabla P_k \phi_0\|_{L_x^2(\mathbb{R}^2)}^2 \right)^{\frac{1}{2}}, \\ &\lesssim \sqrt{p} |I|^{\frac{1}{q}} \|\phi_0\|_{\dot{H}_x^1(\mathbb{R}^2)}, \end{aligned}$$

and the desired result follows from Lemma 3.6 by writing

$$\{\|\partial u^{(0)}\|_{L_t^q L_x^r(I \times \mathbb{R}^2)} < \infty\} = \bigcup_{\ell=1}^{\infty} \{\|\partial u^{(0)}\|_{L_t^q L_x^r(I \times \mathbb{R}^2)} \leq \ell\}.$$

□

#### 4. The iteration scheme

To bound the iterates, we will employ the energy estimates for the wave equation, namely

$$\begin{aligned} &\|u^{(n)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(n)}\|_{L_t^\infty L_x^2} \\ &\lesssim \|u^{(0)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(0)}\|_{L_t^\infty L_x^2} + \|(\partial u^{(n-1)})^2\|_{L_t^1 L_x^2} \\ &= \|u^{(0)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(0)}\|_{L_t^\infty L_x^2} + \|\partial u^{(n-1)}\|_{L_t^2 L_x^4}^2. \end{aligned}$$

Hence, it suffices to obtain bounds for the term

$$\|\partial u^{(n-1)}\|_{L_t^2 L_x^4}.$$

To do this, we will perform an analysis based on a precise representation of the iterates, namely as a sum with products of Rademacher random variables as the coefficients.

In the sequel, to simplify our expression for the iterate expansions, we take  $\phi_1 = 0$  for our initial data. We prove a preliminary version of this representation formula in the next proposition, which we will refine subsequently. In the sequel, we will implicitly regard the indices  $k_i \in \mathbb{Z}^2$  as belonging to  $\mathbb{N}$  via a fixed bijection.

**Proposition 4.1.** *Let  $\{\varepsilon_k\}_{k=1}^\infty$  be the sequence of independent identically distributed Rademacher random variables used in the definition (2.1). We have the representation*

$$\partial u^{(n)} = \sum_{j=1}^{2^n} \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)},$$

where  $k_i \in \mathbb{N}$ , and for any  $n \in \mathbb{N}$ , and  $1 \leq j \leq 2^n$ ,

$$(4.1) \quad G_{k_1}^{(n)} := F_{k_1}, \quad G_{k_1, \dots, k_j}^{(n)} := \sum_{i \in B_j} A_0(G_{k_1, \dots, k_i}^{(n-1)}, G_{k_{i+1}, \dots, k_j}^{(n-1)}), \quad j \geq 2$$

for  $F_{k_1}$  defined in (2.4) and

$$B_j = \left\{ i : 1 \leq i \leq j-1 \text{ if } j \leq 2^{n-1}, \quad j - 2^{n-1} \leq i \leq 2^{n-1} \text{ if } j > 2^{n-1} \right\}.$$

**Remark 4.2.** We point out that  $B_j$  is only defined for  $j \geq 2$  and it is the collection of indices  $i$  for  $1 \leq i \leq 2^{n-1}$  which can contribute to the term with  $j$  Rademacher random variables in the  $n$ -th iterate.

*Proof.* The expression holds for  $n = 0$ . Assume next it holds for  $n - 1$ . We then have from the formula (2.4) that

$$\partial u^{(n)} = \partial u^{(0)} + \sum_{i=1}^{2^{n-1}} \sum_{j=1}^{2^{n-1}} \sum_{k_1, \dots, k_i} \sum_{\ell_1, \dots, \ell_j} \varepsilon_{k_1} \dots \varepsilon_{k_i} \varepsilon_{\ell_1} \dots \varepsilon_{\ell_j} A_0(G_{k_1, \dots, k_i}^{(n-1)}, G_{\ell_1, \dots, \ell_j}^{(n-1)}).$$

Now, since we are summing over all indices, all terms with  $i + j$  many Rademacher random variables may be grouped together since every combination of coefficients appears in front of all of them. Consequently, we group terms with the same number of the Rademacher random variables coefficients, and perform the change of variables  $i + j = j$ . Then with  $B_j$  as

above, we obtain the result since  $B_j$  contains precisely the indices  $i$  which contribute to the  $j$ -th term.  $\square$

Next observe that the expression for  $G_{k_1, \dots, k_j}^{(n)}$  involves  $j - 1$  time integrations and using (4.1) it can be written as

$$(4.2) \quad G_{k_1, \dots, k_j}^{(n)} \sim \underbrace{\int \dots \int}_{j-1 \text{ times}} \tilde{G}_{k_1, \dots, k_j}^{(n)},$$

where

$$(4.3) \quad \begin{aligned} \tilde{G}_{k_1}^{(0)} &:= F_{k_1}, \\ \tilde{G}_{k_1, \dots, k_j}^{(n)} &:= \sum_{i \in B_j} M_{01}(\tilde{G}_{k_1, \dots, k_i}^{(n-1)} \cdot \tilde{G}_{k_{i+1}, \dots, k_j}^{(n-1)}), \quad n \geq 1, \end{aligned}$$

with  $M_{01}$  as given by (3.1). This can be seen readily by induction and the formulas above.

We will now describe the structure of the terms in  $G_{k_1, \dots, k_j}^{(n)}$  more precisely. We observe that these terms involve time integrals with different iterative structures. In the following proposition, we will establish that these different contributions are in bijection with a collection of full binary trees. We recall a full binary tree is a tree where each node (also called a vertex) has either no children or exactly two children. If a node does not have a child, then it is called a leaf (or a terminal vertex). If a node has a child, it is called an internal node. We define the height of a binary tree to be the number of edges between the root and the furthest leaf.

In the sequel for a binary tree  $\tau$  with  $j$  leaves it will be useful to set notation  $\tau = \tau_i \cup \tau_{j-i}$  where  $\tau_i$  and  $\tau_{j-i}$  are the unique trees with  $i$  and  $j - i$  leaves respectively such that the root of  $\tau$  has  $\tau_i$  and  $\tau_{j-i}$  as left and right children.

**Proposition 4.3.** *Let  $n \geq 0$  and  $1 \leq j \leq 2^n$ . Let  $\mathcal{T}$  denote the collection of full binary trees. Then there is an injective map from the terms appearing in  $G_{k_1, \dots, k_j}^{(n)}$ , mapping terms in (4.1) to trees with  $j$  leaves and  $j - 1$  internal nodes. We denote the image of the map by  $\mathcal{T}_j$  and we will write*

$$G_{k_1, \dots, k_j}^{(n)} := \sum_{\tau \in \mathcal{T}_j} G_{k_1, \dots, k_j}^{(n), \tau}.$$

956 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

**Remark 4.4.** Analogously to (4.2), we will use the notation

$$G_{k_1, \dots, k_j}^{(n), \tau} = \underbrace{\int \dots \int}_{j-1 \text{ times}} \tilde{G}_{k_1, \dots, k_j}^{(n), \tau}.$$

**Remark 4.5.** Before the proof, we provide some examples. One example arises from

$$A_0(\partial u^{(0)}, A_0(\partial u^{(0)}, \dots A_0(\partial u^{(0)}, \partial u^{(0)}))),$$

and in this case, we have a contribution of

$$G_{k_1, \dots, k_j}^{(n), \tau} = \underbrace{\int_0^{t_0=t} \int_0^{t_1} \dots \int_0^{t_{j-2}}}_{j-1 \text{ times}} \tilde{G}_{k_1, \dots, k_j}^{(n), \tau} dt_1 \dots dt_{j-1}.$$

This maps to a full binary tree with height  $j - 1$ , where each left node is a leaf. Another example arises from the contribution when  $j = 2^n$ , and

$$G_{k_1, \dots, k_j}^{(n), \tau} = \int_0^t \left( \int_0^{t_1} \dots \left( \int_0^{t_{n-2}} \left( \int_0^{t_{n-1}} \tilde{G}_{k_1, \dots, k_j}^{(n), \tau} dt_n \right)^2 dt_{n-1} \right)^2 \dots dt_2 \right)^2 dt_1.$$

This maps to a full binary tree with height  $n$ , and with leaves appearing only at the final level.

*Proof of Proposition 4.3.* We define the map inductively. Let  $n \geq 0$ , and  $j = 1$ , then  $G_{k_1}^{(n)} = F_{k_1}$ , so there is no integral, and we map this term to a single node. For  $j \geq 2$ , we place a node whenever there is an appearance of the integral operator  $A_0$ , with the left and right children of the node corresponding to the images of the left and right terms in the bilinear operator. Working out one more example for  $n = 1$  and  $j = 2$ , we have

$$G_{k_1, k_2}^{(1)} = A_0(G_{k_1}^{(0)}, G_{k_2}^{(0)}),$$

which would map to a binary tree with a single internal node, and two leaves. We now proceed with the induction.



Let now  $n \geq 1$ , and  $2 \leq j \leq 2^n$ , then by definition, we have

$$(4.4) \quad G_{k_1, \dots, k_j}^{(n)} = \sum_{i \in B_j} A_0(G_{k_1, \dots, k_i}^{(n-1)}, G_{k_{i+1}, \dots, k_j}^{(n-1)}).$$

We take the following formula as the inductive hypothesis

$$G_{k_1, \dots, k_j}^{(n)} = \sum_{\tau \in \mathcal{T}_j} G_{k_1, \dots, k_j}^{(n), \tau}.$$

By (4.4) and the inductive hypothesis we have

$$\begin{aligned} G_{k_1, \dots, k_j}^{(n+1)} &= \sum_{i \in B_j} A_0(G_{k_1, \dots, k_i}^{(n)}, G_{k_{i+1}, \dots, k_j}^{(n)}) \\ &= \sum_{i \in B_j} \sum_{\tau_i \in \mathcal{T}_i} \sum_{\tau_{j-i} \in \mathcal{T}_{j-i}} A_0(G_{k_1, \dots, k_i}^{(n), \tau_i}, G_{k_{i+1}, \dots, k_j}^{(n), \tau_{j-i}}). \end{aligned}$$

Since  $A_0$  gives an integral, we get a collection of trees where the left child comes from trees in  $\mathcal{T}_i$ , and the right child comes from trees in  $\mathcal{T}_{j-i}$  giving a tree with  $i + (j - i) = j$  leaves, and  $i - 1 + (j - i - 1) = j - 1$  internal nodes. Since the decomposition into two children trees is unique, this map is injective.  $\square$

**Corollary 4.6.** *We have the representation*

$$(4.5) \quad G_{k_1, \dots, k_j}^{(n), \tau} = \underbrace{\int_0^{t_0=t} \dots \int_0^{t_{j-2}}}_{j-1 \text{ times}} \tilde{G}_{k_1, \dots, k_j}^{(n), \tau} dt_1 \dots dt_{j-1},$$

with

$$(4.6) \quad \begin{aligned} \tilde{G}_{k_1}^{(0), \tau} &:= F_{k_1}, \\ \tilde{G}_{k_1, \dots, k_j}^{(n), \tau} &:= M_{01}(\tilde{G}_{k_1, \dots, k_i}^{(n-1), \tau_i} \cdot \tilde{G}_{k_{i+1}, \dots, k_j}^{(n-1), \tau_{j-i}}), \end{aligned}$$

where the notation  $\tau_i$  and  $\tau_{j-i}$  means  $\tau_i \in \mathcal{T}_i$  and  $\tau_{j-i} \in \mathcal{T}_{j-i}$  for some trees  $\tau_i$  and  $\tau_{j-i}$  with  $i$  and  $j - i$  leaves respectively.

Furthermore, we observe the support of  $\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}$  is contained in the ball of radius  $j$ .

Let  $I_\tau(t)$  denote the  $j - 1$  iterated time integral which arises in (4.5), that is,  $I_\tau(t)$  is the number which we obtain by replacing  $\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}$  by 1 in (4.5) and carrying out the time integration.

958 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

**Lemma 4.7.** *Let  $n \geq 0$  and let  $1 \leq j \leq 2^n$ . For every  $\tau \in \mathcal{T}_j$ , we have*

$$I_\tau(t) = \frac{t^{j-1}}{C_\tau},$$

*and furthermore, we have the recurrence relation*

$$(4.7) \quad \begin{aligned} C_{\tau_1} &= 1, \quad j = 1, \\ C_\tau &= (j-1) C_{\tau_i} C_{\tau_{j-i}}, \quad \tau \neq \tau_1, \quad j \geq 2. \end{aligned}$$

*where  $\tau_1$  denotes the tree with only one node.*

*Proof.* We argue again by induction. The case  $n = 0$  is clear, and by definition, for  $\tau = \tau_i \cup \tau_{j-i}$  we can represent  $I_\tau$  as

$$I_\tau(t) = \int_0^t dt_1 I_{\tau_i}(t_1) I_{\tau_{j-i}}(t_1),$$

and the result follows from the definition of  $C_\tau$  and integration.  $\square$

It will be useful in the sequel to introduce the notation

$$C_{\tau,j}^* = \inf_{\tau \in \mathcal{T}_j} C_\tau,$$

and we note that the upper bound

$$(4.8) \quad C_{\tau,2^n}^* \leq \prod_{k=1}^n (2^k - 1)^{2^{n-k}}$$

follows from considering the tree of height  $n$ .

The next proposition now simply follows from Proposition 4.3 and Lemma 4.7

**Proposition 4.8.** *Let  $n \geq 1$  and  $2 \leq j \leq 2^n$ . Then*

$$\|G_{k_1, \dots, k_j}^{(n)}\|_{L_t^\infty L_x^4} \leq \sum_{\tau \in \mathcal{T}_j} \frac{|I|^{j-1}}{C_\tau} \|\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}\|_{L_t^\infty L_x^4}.$$

We next turn to establishing a suitable bound for

$$(4.9) \quad \|\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}\|_{L_t^\infty L_x^4}.$$

**Proposition 4.9.** *Let  $n \geq 1$ , and  $2 \leq j \leq 2^n$ . There exists  $C > 0$  such that for any  $\tau \in \mathcal{T}_j$ , we have*

$$\|\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}\|_{L_t^\infty L_x^4} \leq C^{\frac{j-1}{2}} \cdot \sqrt{C_\tau} \prod_{i=1}^j \|P_{k_i} \phi_0\|_{\dot{H}_x^1},$$

where  $C_\tau$  is given in (4.7).

*Proof.* We prove this by induction on  $n$ . For  $n = 1$ , by definition, Lemma 3.1, Lemma 3.3 and Hölder, we have

$$\begin{aligned} \|\tilde{G}_{k_1, k_2}^{(1), \tau}\|_{L_x^4} &= \|M_{01}(F_{k_1} \cdot F_{k_2})\|_{L_x^4} \\ &\leq C_0 \sqrt{2} \pi^{\frac{1}{4}} \|M_{01}(F_{k_1} \cdot F_{k_2})\|_{L_x^2} \\ &\leq C_0 \sqrt{2} \pi^{\frac{1}{4}} \|F_{k_1}\|_{L_x^4} \|F_{k_2}\|_{L_x^4} \\ &\leq C_0^3 \sqrt{2} \pi^{\frac{3}{4}} \|F_{k_1}\|_{L_x^2} \|F_{k_2}\|_{L_x^2}. \end{aligned}$$

For general  $n$ , we similarly have

$$\begin{aligned} \|\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}\|_{L_x^4} &\leq \|M_{01}(\tilde{G}_{k_1, \dots, k_i}^{(n-1), \tau_i} \cdot \tilde{G}_{k_{i+1}, \dots, k_j}^{(n-1), \tau_{j-i}})\|_{L_x^4} \\ &\leq C_0 \pi^{\frac{1}{4}} j^{\frac{1}{2}} \|M_{01}(\tilde{G}_{k_1, \dots, k_i}^{(n-1), \tau_i} \cdot \tilde{G}_{k_{i+1}, \dots, k_j}^{(n-1), \tau_{j-i}})\|_{L_x^2} \\ &\leq C_0 \pi^{\frac{1}{4}} j^{\frac{1}{2}} \|\tilde{G}_{k_1, \dots, k_i}^{(n-1), \tau_i}\|_{L_t^\infty L_x^4} \cdot \|\tilde{G}_{k_{i+1}, \dots, k_j}^{(n-1), \tau_{j-i}}\|_{L_x^4}. \end{aligned}$$

We now let  $C > 0$  be such that

$$(4.10) \quad C_0^2 \pi^{\frac{1}{2}} j \leq C(j-1)$$

for all  $j \geq 2$ . Hence, using the inductive hypothesis and (4.10), we obtain

$$\begin{aligned} \|\tilde{G}_{k_1, \dots, k_j}^{(n), \tau}\|_{L_x^4} &\leq C_0 \pi^{\frac{1}{4}} j^{\frac{1}{2}} C^{\frac{i-1}{2}} \sqrt{C_{\tau_i}} \\ &\quad \times \prod_{\ell=1}^i \|P_{k_\ell} \phi_0\|_{\dot{H}_x^1} C^{\frac{j-i-1}{2}} \sqrt{C_{\tau_{j-i}}} \prod_{\ell=i+1}^j \|P_{k_\ell} \phi_0\|_{\dot{H}_x^1} \\ &\leq C^{\frac{j-2}{2}} \sqrt{C(j-1)C_{\tau_i}C_{\tau_{j-i}}} \prod_{\ell=i}^j \|P_{k_\ell} \phi_0\|_{\dot{H}_x^1}, \end{aligned}$$

which then yields the result by (4.7).  $\square$

**Remark 4.10.** The factor from Bernstein’s inequality will be  $C_0 j^{\frac{3}{4}}$  in dimension three and  $C_0 j$  in dimension four, in which case we will obtain a modified bound for Proposition 4.9, specifically the power of  $C_\tau$  will be  $3/4$  in dimension three and 1 in dimension four. Since Proposition 4.8 already has one power of  $C_\tau$  in the denominator, dimensions two, three and four are the dimensions that will not produce a positive power of  $C_\tau$  in Proposition 4.11 below.

We conclude this section by stating the required  $L_t^2 L_x^4$  bounds we will rely on in the proof of the main theorem.

**Proposition 4.11.** *For any  $n \geq 0$ ,  $1 \leq j \leq 2^n$  and  $G_{k_1, \dots, k_j}^{(n)}$  as above, we have*

$$(4.11) \quad \|G_{k_1, \dots, k_j}^{(n)}\|_{L_t^2 L_x^4} \leq |I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}} \frac{1}{\sqrt{C_{\tau, j}^*}} \prod_{i=1}^j \|P_{k_i} \phi_0\|_{\dot{H}_x^1}.$$

*Proof.* For  $j = 1$ , this holds by Proposition 3.7, while for  $j \geq 2$ , this result is a summary of the previous bounds, together with the fact that the number of full binary trees with  $j$  leaves is given by the  $j$ -th Catalan number, which is exponential in  $j$ , and hence can be absorbed into the  $C^{\frac{j-1}{2}}$  factor.  $\square$

## 5. Proof of main probabilistic bounds

By the discussion at the beginning of Section 4 and in view of Proposition 4.1, the goal of this section is to establish our main probabilistic bounds for the expression

$$\left\| \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_t^2 L_x^4} \right\|_{L_\omega^p}.$$

This will suffice for establishing bounds on the iterates.

Provided  $p \geq 4$ , we can use Minkowski’s inequality to bring the  $L_\omega^p$  norm inside, and hence we first consider

$$\left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_\omega^p}.$$

We group the summation over  $k_1, \dots, k_j$  based on how many distinct indices  $k_i$  appear, ranging from  $r = 1, \dots, j$ . So for example, when  $r = j$ , all the indices are different, while if  $r = 1$ , then all the indices are the same, and the product of the random variables reduces to  $\varepsilon_k^j$ . Moreover, we observe that for each  $r$ , we ask in how many ways we can distribute  $j$  (labeled) indices into  $r$  (unlabeled) groups. Stirling numbers of the second kind provide an answer to this and are labeled by  $S(j, r)$  (more on  $S(j, r)$  below). We let  $P_{j,r}$  denote the collection of such distributions. An element of  $P_{j,r}$  can be identified with  $\vec{k}$  to denote a vector of length  $j$ , which has  $r$  distinct indices  $k_i$  appearing  $\alpha_i$  times so that  $\alpha_1 + \alpha_2 + \dots + \alpha_r = j$ .

Ultimately we need to estimate

$$(5.1) \quad \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_\omega^p} \leq \sum_{r=1}^j \sum_{\vec{k} \in P_{j,r}} \left\| \sum_{k_1, \dots, k_r, k_i \neq k_\ell} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \right\|_{L_\omega^p}.$$

First we consider one of the terms

$$(5.2) \quad \left\| \sum_{k_1, \dots, k_r, k_i \neq k_\ell} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \right\|_{L_\omega^p}.$$

We would like to apply Lemma 3.5 to estimate this expression. However, we are not quite yet in a suitable context since the same random variable may appear in more than one summand. Thus we follow an argument used in [22].

We fix a large  $N \in \mathbb{N}$  and we let  $\mathbb{N}_N := \{1, \dots, N\}$ . We will estimate

$$\sum_{k_1, \dots, k_r, k_i \neq k_\ell, k_i \in \mathbb{N}_N} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)}$$

uniformly in  $N$  which will enable us to conclude the desired bound for (5.2).

We use the identity

$$(5.3) \quad 1 = \frac{1}{r^{N-r}} \sum_{I_1 \sqcup \dots \sqcup I_r = \mathbb{N}_N} \mathbf{1}(k_1 \in I_1) \dots \mathbf{1}(k_r \in I_r),$$

where  $I \sqcup J$  denote a disjoint union, and the sum is over all such disjoint unions of  $r$ -many arbitrary subsets  $I_1, \dots, I_r$  of  $\mathbb{N}_N$ . Since the summand is

zero if any of the  $I_i = \emptyset$ , we will assume that the  $I_i$  are non-empty. We note that this identity holds for all tuples  $(k_1, \dots, k_r)$  with  $k_i \in \mathbb{N}_N$  such that  $k_i \neq k_\ell$  for  $i \neq \ell$ .

The normalization factor in (5.3) is the number of distinct disjoint unions of  $\mathbb{N}_N$  a fixed tuple  $(k_1, \dots, k_r)$  with  $k_i \neq k_\ell$  for  $i \neq \ell$ , may appear in. We compute it as follows: we need to allocate the points

$$\{1, \dots, N\} \setminus \{k_1, \dots, k_r\}$$

to sets  $I_1 \ni k_1, \dots, I_r \ni k_r$ . There are  $r$ -many options for each of the remaining  $N - r$  points, which yields the normalization factor.

Consequently, using (5.3) we can write

$$\begin{aligned} & \sum_{k_1, \dots, k_r, k_i \neq k_\ell, |k_i| \leq N} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \\ &= \frac{1}{r^{N-r}} \sum_{I_1 \sqcup \dots \sqcup I_r = \mathbb{N}_N} \sum_{k_1 \in I_1, \dots, k_r \in I_r} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)}. \end{aligned}$$

We first tackle the case where each  $\alpha_i$  is odd. We may now use Lemma 3.5 and we define

$$b_{k_1} = \sum_{k_2 \in I_2, \dots, k_r \in I_r} \varepsilon_{k_2}^{\alpha_2} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)}.$$

and we write

$$\left\| \sum_{k_1 \in I_1, \dots, k_r \in I_r} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \right\|_{L_\omega^p} = \left\| \sum_{k_1 \in I_1} \varepsilon_{k_1}^{\alpha_1} b_{k_1} \right\|_{L_\omega^p}.$$

By construction, it is now the case that the family  $\{b_{k_1}\}$  is independent of the family  $\{\varepsilon_{k_1}\}$ , and applying Lemma 3.5, we obtain

$$\begin{aligned} \left\| \sum_{k_1 \in I_1} \varepsilon_{k_1}^{\alpha_1} b_{k_1} \right\|_{L_\omega^p} &\leq C\sqrt{p} \left\| \left( \sum_{k_1 \in I_1} |b_{k_1}|^2 \right)^{\frac{1}{2}} \right\|_{L_\omega^p} \\ &\leq C\sqrt{p} \left( \sum_{k_1 \in I_1} \|b_{k_1}\|_{L_\omega^p}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

We iterate this process and eventually we get

$$\left\| \sum_{k_1 \in I_1, \dots, k_r \in I_r} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \right\|_{L_\omega^p} \leq C^r p^{\frac{r}{2}} \left( \sum_{k_1 \in I_1, \dots, k_r \in I_r} |G_{\vec{k}}^{(n)}|^2 \right)^{\frac{1}{2}}.$$

In the general case, where the  $\alpha_i$  are not all odd, we order the indices and apply the triangle inequality for the even indices. More precisely, if there are even powers  $\alpha_i$ , then  $\varepsilon_{k_i}^{\alpha_i} = 1$ , and in that case, we use the triangle inequality at that step in the estimates, for instance

$$\left\| \sum_{k_1 \in I_1} \varepsilon_{k_1}^{\alpha_1} b_{k_1} \right\|_{L_\omega^p} \leq \sum_{k_1 \in I_1} \|b_{k_1}\|_{L_\omega^p}.$$

Thus, reordering the  $k_i$  and letting  $r_o$  denote the number of odd  $\alpha_i$ , we obtain

$$\left\| \sum_{k_1 \in I_1} \varepsilon_{k_1}^{\alpha_1} b_{k_1} \right\|_{L_\omega^p} \leq \sum_{k_i \in I_i, \alpha_i \text{ even}} C^{r_o} p^{\frac{r_o}{2}} \left( \sum_{k_i \in I_i, \alpha_i \text{ odd}} |G_{\vec{k}}^{(n)}|^2 \right)^{\frac{1}{2}}.$$

We put everything together and we have

$$\begin{aligned} & \sum_{r=1}^j \sum_{\vec{k} \in P_{j,r}} \left\| \sum_{k_1, \dots, k_r, k_i \neq k_\ell, |k_i| \leq N} \varepsilon_{k_1}^{\alpha_1} \dots \varepsilon_{k_r}^{\alpha_r} G_{\vec{k}}^{(n)} \right\|_{L_\omega^p} \\ & \leq \sum_{r=1}^j \sum_{\vec{k} \in P_{j,r}} \frac{1}{r^{N-r}} \sum_{I_1 \sqcup \dots \sqcup I_r = \mathbb{N}_N} \sum_{k_i \in I_i, \alpha_i \text{ even}} C^{r_o} p^{\frac{r_o}{2}} \left( \sum_{k_i \in I_i, \alpha_i \text{ odd}} |G_{\vec{k}}^{(n)}|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Now we estimate the expression

$$\frac{1}{r^{N-r}} \sum_{I_1 \sqcup \dots \sqcup I_r = \mathbb{N}_N} 1.$$

The sum will count the number of surjective functions on  $N$  items into  $r$  items, and by inclusion exclusion this is equal to

$$r!S(N, r),$$

964 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

where  $S(N, r)$  is the Stirling number of the second kind, given by the formula

$$(5.4) \quad S(N, r) = \frac{1}{r!} \sum_{j=0}^r (-1)^{r-j} \binom{r}{j} j^N.$$

We can use the trivial bound

$$r!S(N, r) \leq r^N,$$

which is the total number of functions on  $N$  items into  $r$  items, and hence we obtain that

$$(5.5) \quad \frac{1}{r^{N-r}} \sum_{I_1 \sqcup \dots \sqcup I_r = \mathbb{N}_N} 1 = \frac{r!S(N, r)}{r^{N-r}} \leq r^r.$$

Thus we obtain a uniform bound in  $N$ , which yields the desired bound

$$(5.6) \quad \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_\omega^p} \leq \sum_{r=1}^j \sum_{\vec{k} \in P_{j,r}} \sum_{k_i, \alpha_i \text{ even}} C^{r_0} p^{\frac{r_0}{2}} r^r \left( \sum_{k_i, \alpha_i \text{ odd}} |G_{\vec{k}}^{(n)}|^2 \right)^{\frac{1}{2}}.$$

We are now ready to establish the relevant bound for (5.1) in the following proposition.

**Proposition 5.1.** *Let  $p \geq 4$ , and let  $G_{k_1, \dots, k_j}^{(n)}$  be given as above. Then*

$$\left\| \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_x^2 L_x^4} \right\|_{L_\omega^p} \leq C^j \|\phi_0\|_{\dot{H}_x^1}^j |I|^{j-\frac{1}{2}} p^{\frac{j}{2}} \frac{j!}{\sqrt{C_{\tau,j}^*}}$$

for a constant  $C$  independent of  $n$ .

*Proof.* In this proof, we will use  $C$  to denote explicit constants whose values do not depend on the various parameters. We allow the value of  $C$  to change line to line.



By Minkowski's inequality we first have

$$\left\| \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_t^2 L_x^4} \right\|_{L_\omega^p} \leq \left\| \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_\omega^p} \right\|_{L_t^2 L_x^4},$$

so we can take the  $L_t^2 L_x^4$  norm of (5.6) and use (4.11) to obtain

$$\begin{aligned} & \sum_{r=1}^j \sum_{\vec{k} \in P_{j,r}} \sum_{k_i, \alpha_i \text{ even}} C^{r_o} p^{\frac{r_o}{2}} r^r \left( \sum_{k_i, \alpha_i \text{ odd}} \|G_{\vec{k}}^{(n)}\|_{L_t^2 L_x^4}^2 \right)^{\frac{1}{2}} \\ & \leq \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} \sum_{r=1}^j C^r p^{\frac{r}{2}} r^r \sum_{\vec{k} \in P_{j,r}} \sum_{k_i, \alpha_i \text{ even}} \left( \sum_{k_i, \alpha_i \text{ odd}} \left( \prod_{i=1}^r \|P_{k_i} \phi_0\|_{\dot{H}_x^1}^{\alpha_i} \right)^2 \right)^{\frac{1}{2}} \\ & \leq \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} C^j p^{\frac{j}{2}} \sum_{r=1}^j r^r \sum_{\vec{k} \in P_{j,r}} \sum_{k_i, \alpha_i \text{ even}} \left( \sum_{k_i, \alpha_i \text{ odd}} \prod_{i=1}^r \|P_{k_i} \phi_0\|_{\dot{H}_x^1}^{2\alpha_i} \right)^{\frac{1}{2}}, \end{aligned}$$

where we used  $r_o \leq r \leq j$ . Next we observe that

$$\sum_{k_i, \alpha_i \text{ even}} \left( \sum_{k_i, \alpha_i \text{ odd}} \prod_{i=1}^r \|P_{k_i} \phi_0\|_{\dot{H}_x^1}^{2\alpha_i} \right)^{\frac{1}{2}} \leq \|\phi_0\|_{\dot{H}_x^1}^j.$$

We are then left with estimating

$$\|\phi_0\|_{\dot{H}_x^1}^j \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} C^j p^{\frac{j}{2}} \sum_{r=1}^j r^r \sum_{\vec{k} \in P_{j,r}} 1.$$

Since there are  $S(j, r)$  many vectors in  $P_{j,r}$  we use a refined upper bound for the Stirling numbers of the second kind (c.f. [54, Theorem 3]):

$$S(j, r) \leq \frac{1}{2} \binom{j}{r} r^{j-r}.$$

Together with the upper bound for the binomial coefficients

$$\binom{j}{r} \leq \left( \frac{ej}{r} \right)^r,$$

966 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

this yields

$$\begin{aligned} & \|\phi_0\|_{\dot{H}_x^1}^j \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} C^j p^{\frac{j}{2}} \sum_{r=1}^j \frac{1}{2} \left(\frac{ej}{r}\right)^r r^j \\ &= \|\phi_0\|_{\dot{H}_x^1}^j \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} C^j p^{\frac{j}{2}} \sum_{r=1}^j \frac{1}{2} e^r j^r r^{j-r} \\ &\leq \|\phi_0\|_{\dot{H}_x^1}^j \frac{|I|^{j-\frac{1}{2}} C^{\frac{j-1}{2}}}{\sqrt{C_{\tau,j}^*}} C^j p^{\frac{j}{2}} j^j \sum_{r=1}^j \frac{1}{2} e^r. \end{aligned}$$

Ultimately, we obtain that there exists some constant  $C$  (possibly different from above, but still uniform in  $n$ ) such that

$$(5.7) \quad \left\| \left\| \sum_{k_1, \dots, k_j} \varepsilon_{k_1} \dots \varepsilon_{k_j} G_{k_1, \dots, k_j}^{(n)} \right\|_{L_t^2 L_x^4} \right\|_{L_\omega^p} \leq C^j \|\phi_0\|_{\dot{H}_x^1}^j |I|^{j-\frac{1}{2}} p^{\frac{j}{2}} \frac{j!}{\sqrt{C_{\tau,j}^*}}.$$

This yields the desired bound and concludes the proof.  $\square$

**Corollary 5.2.** *Let  $n \geq 0$ . Then*

$$\left\| \|\partial u^{(n)}\|_{L_t^2 L_x^4} \right\|_{L_\omega^p} \leq \sum_{j=1}^{2^n} C^j \|\phi_0\|_{\dot{H}_x^1}^j |I|^{j-\frac{1}{2}} p^{\frac{j}{2}} j!$$

for a constant  $C$  independent of  $n$ .

*Proof.* This follows by Proposition [4.1](#) and the trivial lower bound  $C_{\tau,j}^* \geq 1$ .  $\square$

**Remark 5.3.** We note that using the upper bound in [\(4.8\)](#), we have that

$$C_{\tau, 2^n}^* \leq \prod_{k=1}^n 2^{k2^{n-k}}.$$

Now we can compute exactly that

$$\sum_{k=1}^n k 2^{n-k} = 2^{n+1} - n - 2,$$

hence we see the factor of  $C_{\tau,j}^*$  in our bound (5.7) does not suffice to cancel the factorial growth.

## 6. Proof of Theorem 2.1

The main theorem will follow from the following result and Lemma 3.6.

**Proposition 6.1.** *There exist constants  $C > 0$  and  $\delta > 0$  such that for every  $n \in \mathbb{N}$ , and every interval  $I$  with  $|I| < \delta$  the following holds*

$$\left\| \left\| \partial u^{(n)} \right\|_{L_t^2 L_x^4(I \times \mathbb{R}^2)} \right\|_{L_\omega^p} \leq C p^{\frac{2n}{2}} \|\phi_0\|_{\dot{H}_x^1} |I|^{\frac{1}{2}} (2^n)!$$

*Proof.* By Corollary 5.2

$$\left\| \left\| \partial u^{(n)} \right\|_{L_t^2 L_x^4} \right\|_{L_\omega^p} \leq \sum_{j=1}^{2^n} C^j \|\phi_0\|_{\dot{H}_x^1}^j |I|^{j-\frac{1}{2}} p^{\frac{j}{2}} j!.$$

We rewrite this bound as

$$\begin{aligned} & C \|\phi_0\|_{\dot{H}_x^1} |I|^{\frac{1}{2}} \sum_{j=1}^{2^n} C^{j-1} \|\phi_0\|_{\dot{H}_x^1}^{j-1} |I|^{j-1} p^{\frac{j}{2}} j! \\ & \leq C \|\phi_0\|_{\dot{H}_x^1} |I|^{\frac{1}{2}} p^{\frac{2n}{2}} (2^n)! \sum_{j=0}^{2^n-1} C^j \|\phi_0\|_{\dot{H}_x^1}^j |I|^j, \end{aligned}$$

and we note that the sum in  $j$  is bounded by one, say, provided we choose  $I \subseteq \mathbb{R}$  such that

$$C \|\phi_0\|_{\dot{H}_x^1} |I| < \frac{1}{2}.$$

□

**Corollary 6.2.** *There exist  $C, c > 0$ , independent of  $n$ , such that for any  $I \subseteq \mathbb{R}$  with  $|I|$  sufficiently small, we have*

$$\mathbb{P}\left\{\left\|\partial u^{(n)}\right\|_{L_t^2 L_x^4(I \times \mathbb{R}^2)} > \lambda\right\} \leq C \exp\left(-\frac{c \lambda^{\frac{1}{2^n-1}}}{(2^n \|\phi_0\|_{\dot{H}_x^1} |I|^{\frac{1}{2}})^{\frac{1}{2^n-1}}}\right).$$

*In particular, for any fixed  $n$ , we have*

$$\mathbb{P}\left\{\left\|\partial u^{(n)}\right\|_{L_t^2 L_x^4(I \times \mathbb{R}^2)} < \infty\right\} = 1.$$

968 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

*Proof.* This is an application of Lemma 3.6 to the result of Proposition 6.1 with  $k = 2^n$ ,  $\alpha = 1$ .  $\square$

The proof of our main theorem now follows readily from these estimates.

*Proof of Theorem 2.1.* We fix  $(\phi_0, \phi_1)$  and  $T > 0$  so that  $T$  is sufficiently small in the sense required for Corollary 6.2. By the energy estimates, we recall that we have

$$(6.1) \quad \begin{aligned} & \|u^{(n)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(n)}\|_{L_t^\infty L_x^2} \\ & \lesssim \|u^{(0)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(0)}\|_{L_t^\infty L_x^2} + \|\partial u^{(n-1)}\|_{L^2 L_x^4}^2. \end{aligned}$$

Since  $\phi^\omega$  almost surely in  $\dot{H}^1(\mathbb{R}^2) \times L_x^2(\mathbb{R}^2)$ , we let  $\Sigma$  be such that for  $\omega \in \Sigma$ ,

$$\|\phi_0^\omega\|_{L_x^\infty \dot{H}_x^1} + \|\phi_1^\omega\|_{L_x^\infty L_x^2} < \infty,$$

which then provides a bound for

$$\|u^{(0)}\|_{L_t^\infty \dot{H}_x^1} + \|\partial_t u^{(0)}\|_{L_t^\infty L_x^2}.$$

Now, for  $n \geq 0$ , we let

$$\Sigma_n = \{\omega : \|\partial u^{(n)}\|_{L_t^2 L_x^4(I \times \mathbb{R}^2)} < \infty\},$$

then  $\Sigma_n$  has full measure by Proposition 3.7 and Corollary 6.2. We then set

$$\Sigma_T := \Sigma \cap \bigcap_{n=0}^{\infty} \Sigma_n,$$

and we observe that  $\mathbb{P}(\Sigma_T) = 1$  since it is the countable intersection of sets of full measure.

Now, for any  $\omega \in \Sigma_T$ , and any  $n \geq 0$ , we have by (6.1) that

$$(u^{(n)}, \partial_t u^{(n)}) \in L_t^\infty([0, T]; \dot{H}_x^1(\mathbb{R}^2) \times L_x^2(\mathbb{R}^2)).$$

The continuity follows from the definition of the iterates. This completes the proof.  $\square$

## References

- [1] A. Benedek and R. Panzone, *The space  $L^p$ , with mixed norm*, Duke Math. J. **28** (1961), 301–324.

- [2] A. Bényi, T. Oh, and O. Pocovnicu, *Wiener randomization on unbounded domains and an application to almost sure well-posedness of NLS*, Excursions in Harmonic Analysis **4** (2015), 3–25. Appl. Numer. Harmon. Anal., Birkhäuser/Springer, Cham, (2015).
- [3] Á. Bényi, T. Oh, and O. Pocovnicu, *Higher order expansions for the probabilistic local Cauchy theory of the cubic nonlinear Schrödinger equation on  $\mathbb{R}^3$* , Trans. Amer. Math. Soc. Ser. B **6** (2019), 114–160.
- [4] J. Bourgain, *On the Cauchy and invariant measure problem for the periodic Zakharov system*, Duke Math. J. **76** (1994), no. 1, 175–202.
- [5] J. Bourgain, *Periodic nonlinear Schrödinger equation and invariant measures*, Comm. Math. Phys. **166** (1994), no. 1, 1–26.
- [6] J. Bourgain, *Invariant measures for the 2D-defocusing nonlinear Schrödinger equation*, Comm. Math. Phys. **176** (1996), no. 2, 421–445.
- [7] J. Bourgain, *Invariant measures for the Gross-Piatevskii equation*, J. Math. Pures Appl. (9) **76** (1997), no. 8, 649–702.
- [8] N. Bournaveas and T. Candy, *Local well-posedness for the space-time monopole equation in Lorenz gauge*, NoDEA Nonlinear Differential Equations Appl. **19** (2012), no. 1, 67–78.
- [9] N. Burq and N. Tzvetkov, *Invariant measure for a three dimensional nonlinear wave equation*, Int. Math. Res. Not. (2007), no. 22.
- [10] N. Burq and N. Tzvetkov *Random data Cauchy theory for supercritical wave equations. I. Local theory*, Invent. Math. **173** (2008), no. 3, 449–475.
- [11] N. Burq and N. Tzvetkov, *Random data Cauchy theory for supercritical wave equations. II. A global existence result*, Invent. Math. **173** (2008), no. 3, 477–496.
- [12] N. Burq and N. Tzvetkov, *Probabilistic well-posedness for the cubic wave equation*, J. Eur. Math. Soc. **16** (2014), no. 1, 1–30.
- [13] D. Christodoulou, *Global solutions of nonlinear hyperbolic equations for small initial data*, Comm. Pure Appl. Math. **39** (1986), no. 2, 267–282.
- [14] R. Côte, C. E. Kenig, A. Lawrie, and W. Schlag, *Characterization of large energy solutions of the equivariant wave map problem: II*, Amer. J. Math. **137** (2015), no. 1, 209–250.
- [15] M. Czubak, Ph.D. Thesis, University of Texas at Austin (2008).

970 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

- [16] M. Czubak, *Local wellposedness for the  $2+1$ -dimensional monopole equation*, Anal. PDE **3** (2010), no. 2, 151–174.
- [17] M. Czubak and N. Pikula, *Low regularity well-posedness for the 2D Maxwell-Klein-Gordon equation in the Coulomb gauge*, Commun. Pure Appl. Anal. **13** (2014), no. 4, 1669–1683.
- [18] B. Dai and C.-L. Terng, *Bäcklund transformations, Ward solitons, and unitons*, J. Differential Geom. **75** (2007), no. 1, 57–108.
- [19] B. Dai, C.-L. Terng, and K. Uhlenbeck, *On the space-time monopole equation*, in: Surveys in Differential Geometry. Vol. X, pp. 1–30, Int. Press, Somerville, MA (2006).
- [20] P. D’Ancona, D. Foschi, and S. Selberg, *Product estimates for wave-Sobolev spaces in  $2+1$  and  $1+1$  dimensions*, in: Nonlinear Partial Differential Equations and Hyperbolic Wave Phenomena, Vol. 526 of Contemp. Math., pp. 125–150, Amer. Math. Soc., Providence, RI (2010).
- [21] B. Dodson, J. Lührmann, and D. Mendelson, *Almost sure scattering for the 4D energy-critical defocusing nonlinear wave equation with radial data*, Adv. Math. **347** (2019), 619–676.
- [22] L. Erdos and H.-T. Yau, *A Dynamical Approach to Random Matrix Theory*, Vol. 28 of Courant Lecture Notes in Mathematics, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI (2017), ISBN 978-1-4704-3648-3.
- [23] D. Foschi and S. Klainerman, *Bilinear space-time estimates for homogeneous wave equations*, Ann. Sci. École Norm. Sup. (4) **33** (2000), no. 2, 211–274.
- [24] D. S. Freed and K. K. Uhlenbeck, *Instantons and Four-Manifolds*, Vol. 1 of Mathematical Sciences Research Institute Publications, Springer-Verlag, New York (1984), ISBN 0-387-96036-8.
- [25] J. Glimm and A. Jaffe, *Quantum Physics*, Springer-Verlag, New York, second edition (1987), ISBN 0-387-96476-2. A functional integral point of view.
- [26] V. Grigoryan and A. R. Nahmod, *Almost critical well-posedness for nonlinear wave equations with  $Q_{\mu\nu}$  null forms in 2D*, Math. Res. Lett. **21** (2014), no. 2, 313–332.
- [27] V. Grigoryan and A. Tanguay, *Improved well-posedness for the quadratic derivative nonlinear wave equation in 2D*, J. Math. Anal. Appl. **475** (2019), 1578–1595.

- [28] A. Grünrock, *On the wave equation with quadratic nonlinearities in three space dimensions*, J. Hyperbolic Differ. Equ. **8** (2011), no. 1, 1–8.
- [29] S. Klainerman, *Long time behaviour of solutions to nonlinear wave equations*, in: Proceedings of the International Congress of Mathematicians, Vols. 1, 2 (Warsaw, 1983), pp. 1209–1215, PWN, Warsaw (1984).
- [30] S. Klainerman, *The null condition and global existence to nonlinear wave equations*, in: Nonlinear Systems of Partial Differential Equations in Applied Mathematics, Part 1 (Santa Fe, N.M., 1984), Vol. 23 of Lectures in Appl. Math., pp. 293–326, Amer. Math. Soc., Providence, RI (1986).
- [31] S. Klainerman and M. Machedon, *Space-time estimates for null forms and the local existence theorem*, Comm. Pure Appl. Math. **46** (1993), no. 9, 1221–1268.
- [32] S. Klainerman and M. Machedon, *Finite energy solutions of the Yang-Mills equations in  $\mathbf{R}^{3+1}$* , Ann. of Math. (2) **142** (1995), no. 1, 39–119.
- [33] S. Klainerman and M. Machedon, *Smoothing estimates for null forms and applications*, Duke Math. J. **81** (1995), no. 1, 99–133 (1996). A celebration of John F. Nash, Jr.
- [34] S. Klainerman and M. Machedon, *Estimates for null forms and the spaces  $H_{s,\delta}$* , Internat. Math. Res. Notices (1996), no. 17, 853–865.
- [35] S. Klainerman and S. Selberg, *Remark on the optimal regularity for equations of wave maps type*, Comm. Partial Differential Equations **22** (1997), no. 5-6, 901–918.
- [36] S. Klainerman and S. Selberg, *Bilinear estimates and applications to nonlinear wave equations*, Commun. Contemp. Math. **4** (2002), no. 2, 223–295.
- [37] J. Krieger, *Global regularity of wave maps from  $\mathbf{R}^{2+1}$  to  $H^2$ . Small energy*, Comm. Math. Phys. **250** (2004), no. 3, 507–580.
- [38] J. Krieger and J. Lührmann, *Concentration compactness for the critical Maxwell-Klein-Gordon equation*, Ann. PDE **1** (2015), no. 1, Art. 5, 208.
- [39] J. Krieger and W. Schlag, *Concentration Compactness for Critical Wave Maps*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich (2012), ISBN 978-3-03719-106-4.

972 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

- [40] J. Krieger, W. Schlag, and D. Tataru, *Renormalization and blow up for the critical Yang-Mills problem*, Adv. Math. **221** (2009), no. 5, 1445–1521.
- [41] J. Krieger, J. Sterbenz, and D. Tataru, *Global well-posedness for the Maxwell-Klein-Gordon equation in  $4 + 1$  dimensions: small energy*, Duke Math. J. **164** (2015), no. 6, 973–1040.
- [42] J. Krieger and D. Tataru, *Global well-posedness for the Yang-Mills equation in  $4 + 1$  dimensions. Small energy*, Ann. of Math. (2) **185** (2017), no. 3, 831–893.
- [43] J. Lebowitz, H. Rose, and E. Speer, *Statistical mechanics of the nonlinear Schrödinger equation*, J. Statist. Phys. **50** (1988), no. 3-4, 657–687.
- [44] H. Lindblad, *Counterexamples to local existence for semi-linear wave equations*, Amer. J. Math. **118** (1996), no. 1, 1–16.
- [45] J. Lührmann and D. Mendelson, *Random data Cauchy theory for nonlinear wave equations of power-type on  $\mathbb{R}^3$* , Comm. Partial Differential Equations **39** (2014), no. 12, 2262–2283.
- [46] J. Lührmann and D. Mendelson, *On the almost sure global well-posedness of energy sub-critical nonlinear wave equations on  $\mathbb{R}^3$* , New York J. Math. **22** (2016), 209–227.
- [47] V. Moncrief, *Global existence of Maxwell-Klein-Gordon fields in  $(2 + 1)$ -dimensional spacetime*, J. Math. Phys. **21** (1980), no. 8, 2291–2296.
- [48] A. Nahmod, A. Stefanov, and K. Uhlenbeck, *On the well-posedness of the wave map problem in high dimensions*, Comm. Anal. Geom. **11** (2003), no. 1, 49–83.
- [49] S.-J. Oh and D. Tataru, *Global well-posedness and scattering of the  $(4 + 1)$ -dimensional Maxwell-Klein-Gordon equation*, Invent. Math. **205** (2016), no. 3, 781–877.
- [50] T. Oh and O. Pocovnicu, *Probabilistic global well-posedness of the energy-critical defocusing quintic nonlinear wave equation on  $\mathbb{R}^3$* , J. Math. Pures Appl. (9) **105** (2016), no. 3, 342–366.
- [51] H. Pecher, *Low regularity solutions for the  $(2 + 1)$ -dimensional Maxwell-Klein-Gordon equations in temporal gauge*, Commun. Pure Appl. Anal. **15** (2016), no. 6, 2203–2219.



- [52] O. Pocovnicu, *Almost sure global well-posedness for the energy-critical defocusing nonlinear wave equation on  $\mathbb{R}^d$ ,  $d = 4$  and 5*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 8, 2521–2575.
- [53] G. Ponce and T. C. Sideris, *Local regularity of nonlinear wave equations in three space dimensions*, Comm. Partial Differential Equations **18** (1993), no. 1-2, 169–177.
- [54] B. C. Rennie and A. J. Dobson, *On Stirling numbers of the second kind*, J. Combinatorial Theory **7** (1969) 116–121.
- [55] I. Rodnianski and J. Sterbenz, *On the formation of singularities in the critical  $O(3)$   $\sigma$ -model*, Ann. of Math. (2) **172** (2010), no. 1, 187–242.
- [56] J. Shatah and M. Struwe, *The Cauchy problem for wave maps*, Int. Math. Res. Not. (2002), no. 11, 555–571.
- [57] A. de Suzzoni, *Large data low regularity scattering results for the wave equation on the Euclidean space*, Commun. Partial Differ. Equations **38** (2013), no. 1-3, 1–49.
- [58] A. de Suzzoni, *Consequences of the choice of a particular basis of  $L^2(S^3)$  for the cubic wave equation on the sphere and the Euclidean space*, Commun. Pure Appl. Anal. **13** (2014), no. 3, 991–1015.
- [59] J. Sterbenz and D. Tataru, *Regularity of wave-maps in dimension  $2 + 1$* , Comm. Math. Phys. **298** (2010), no. 1, 231–264.
- [60] M. Struwe, *Equivariant wave maps in two space dimensions*, Comm. Pure Appl. Math. **56** (2003), no. 7, 815–823. Dedicated to the memory of Jürgen K. Moser.
- [61] A. J. Tanguay, *New Bilinear Estimates for Quadratic-Derivative Nonlinear Wave Equations in  $2 + 1$  Dimensions*, ProQuest LLC, Ann Arbor, MI (2012), ISBN 978-1267-78809-2. Thesis (Ph.D.)—University of Massachusetts Amherst.
- [62] T. Tao, *Global regularity of wave maps. II. Small energy in two dimensions*, Comm. Math. Phys. **224** (2001), no. 2, 443–544.
- [63] A. Tesfahun, *Almost critical local well-posedness for the space-time monopole equation in Lorenz gauge*, Commun. Contemp. Math. **17** (2015), no. 3, 1450043, 14.
- [64] N. Tzvetkov, *Construction of a Gibbs measure associated to the periodic Benjamin-Ono equation*, Probab. Theory Related Fields **146** (2010), no. 3-4, 481–514.

974 Chanillo, Czubak, Mendelson, Nahmod, and Staffilani

- [65] K. K. Uhlenbeck, *Connections with  $L^p$  bounds on curvature*, Comm. Math. Phys. **83** (1982), no. 1, 31–42.
- [66] K. K. Uhlenbeck, *Removable singularities in Yang-Mills fields*, Comm. Math. Phys. **83** (1982), no. 1, 11–29.
- [67] R. S. Ward, *Soliton solutions in an integrable chiral model in  $2+1$  dimensions*, J. Math. Phys. **29** (1988), no. 2, 386–389.
- [68] R. S. Ward, *Twistors in  $2+1$  dimensions*, J. Math. Phys. **30** (1989), no. 10, 2246–2251.
- [69] D. Wu, *The Cauchy problem of the Ward equation*, J. Funct. Anal. **256** (2009), no. 1, 215–257.
- [70] T. Zhang and D. Fang, *Random data Cauchy theory for the generalized incompressible Navier-Stokes equations*, J. Math. Fluid Mech. **14** (2012), no. 2, 311–324.
- [71] P. E. Zhidkov, *An invariant measure for the nonlinear Schrödinger equation*, Dokl. Akad. Nauk SSSR **317** (1991), no. 3, 543–546.
- [72] P. E. Zhidkov, *On an infinite sequence of invariant measures for the cubic nonlinear Schrödinger equation*, Int. J. Math. Math. Sci. **28** (2001), no. 7, 375–394.
- [73] Y. Zhou, *Local existence with minimal regularity for nonlinear wave equations*, Amer. J. Math. **119** (1997), no. 3, 671–703.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY  
PISCATAWAY, NJ 08854-8019, USA  
*E-mail address:* `chanillo@math.rutgers.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO AT BOULDER  
CAMPUS BOX 395 BOULDER, CO 80309-0395, USA  
*E-mail address:* `czubak@colorado.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO  
5734 S. UNIVERSITY AVENUE, CHICAGO, IL 60637, USA  
*E-mail address:* `dana@math.uchicago.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MASSACHUSETTS AMHERST  
710 NORTH PLEASANT ST. LEDERLE GRADUATE RESEARCH TOWER AMHERST,  
MA 01003-9305, USA  
*E-mail address:* `nahmod@math.umass.edu`

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY  
77 MASSACHUSETTS AVE, ROOM 2-251, CAMBRIDGE, MA 02139-4307, USA  
*E-mail address:* `gigliola@math.mit.edu`

RECEIVED MAY 21, 2018  
ACCEPTED APRIL 1, 2019

