

# ALMOST SURE SCATTERING FOR THE ENERGY CRITICAL NONLINEAR WAVE EQUATION

BJOERN BRINGMANN

ABSTRACT. We study the defocusing energy critical nonlinear wave equation in four dimensions. Our main result proves the stability of the scattering mechanism under random perturbations of the initial data. The random perturbation is defined through a microlocal randomization, which is based on a unit-scale decomposition in physical and frequency space. In contrast to the previous literature, we do not require the spherical symmetry of the perturbation.

The main novelty lies in a wave packet decomposition of the random linear evolution. Through this decomposition, we can adaptively estimate the interaction between the rough and regular components of the solution. Our argument relies on techniques from restriction theory, such as Bourgain’s bush argument and Wolff’s induction on scales.

## CONTENTS

1.	Introduction	1
1.1.	The random data Cauchy problem	2
1.2.	Main result and ideas	4
2.	Notation and preliminaries	8
2.1.	Probability theory	9
2.2.	Harmonic analysis	10
2.3.	Strichartz estimates	13
3.	Probabilistic Strichartz estimates	14
4.	Wave packet decomposition	17
5.	Nonlinear evolution: Local well-posedness, stability theory, and flux estimates	28
5.1.	Local well-posedness and stability theory	28
5.2.	Flux estimates	29
6.	The energy increment and induction on scales	32
	References	41

## 1. INTRODUCTION

We consider the defocusing cubic nonlinear wave equation in four space dimensions, that is,

$$(1) \quad \begin{cases} -\partial_{tt}u + \Delta u = u^3 & (t, x) \in \mathbb{R} \times \mathbb{R}^4, \\ u|_{t=0} = u_0 \in \dot{H}^s(\mathbb{R}^4), \quad \partial_t u|_{t=0} = u_1 \in \dot{H}^{s-1}(\mathbb{R}^4). \end{cases}$$

If  $u$  is a regular solution of (1), then it conserves the energy

$$(2) \quad E[u](t) := \int_{\mathbb{R}^4} \frac{|\nabla u(t, x)|^2}{2} + \frac{(\partial_t u(t, x))^2}{2} + \frac{u(t, x)^4}{4} dx .$$

From the Sobolev embedding  $\dot{H}^1(\mathbb{R}^4) \hookrightarrow L^4(\mathbb{R}^4)$ , it follows that the initial data has finite energy if and only if  $(u_0, u_1) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$ . Thus, we also refer to  $\dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  as the energy space. In addition to the energy conservation law, (1) obeys the scaling symmetry  $u(t, x) \mapsto u_\lambda(t, x) = \lambda u(\lambda t, \lambda x)$ . Since the scaling leaves the energy invariant, the equation is called energy critical. Due to the positive sign in front of the potential term  $u^4$ , we call (1) defocusing. There also exists analogues of (1) with a power-type nonlinearity in any dimension  $d \geq 3$ .

The Cauchy problem for (deterministic) initial data in the energy space is well-understood. We summarize the relevant results in the following theorem.

**Theorem 1.1** (Global well-posedness and scattering [1, 28, 29, 45, 47, 48, 49, 50, 52]).

Let  $(u_0, u_1) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$ . Then, there exists a maximal time interval of existence  $I$  and a unique solution  $u: I \times \mathbb{R}^4 \rightarrow \mathbb{R}$  of (1) such that  $u \in C_t^0 \dot{H}_x^1(I \times \mathbb{R}^4) \cap L_{t, \text{loc}}^3 L_x^6(I \times \mathbb{R}^4)$  and  $\partial_t u \in C_t^0 L_x^2(I \times \mathbb{R}^4)$ . Furthermore, we have that

- (i)  $u$  is global, i.e.,  $I = \mathbb{R}$ .
- (ii)  $u$  obeys a global space-time bound of the form

$$\|u\|_{L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)} \leq C(E[u_0, u_1]) .$$

- (iii)  $u$  scatters to a solution of the linear wave equation. Thus, there exist scattering states  $(u_0^\pm, u_1^\pm) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  s.t.

$$\lim_{t \rightarrow \pm\infty} \|(u(t) - W(t)(u_0^\pm, u_1^\pm), \partial_t u(t) - \partial_t W(t)(u_0^\pm, u_1^\pm))\|_{\dot{H}^1 \times L^2} = 0 .$$

Here,  $W(t)(u_0^\pm, u_1^\pm) = \cos(t|\nabla|)u_0^\pm + (\sin(t|\nabla|)/|\nabla|)u_1^\pm$  denotes the solution to the linear wave equation with initial data  $(u_0^\pm, u_1^\pm)$ .

Global well-posedness and scattering results such as Theorem 1.1 are known for many defocusing dispersive partial differential equations, and hold for the energy critical nonlinear Schrödinger equation [9, 17, 46, 55], the mass-critical nonlinear Schrödinger equation [20, 21, 22, 33, 34], the mass-critical generalized KdV [23], and the  $\dot{H}^{\frac{1}{2}}$ -critical radial nonlinear wave equation [24].

Since Theorem 1.1 provides a complete description of the Cauchy problem with initial data in the energy space, we now seek a similar result for initial data in a rough Sobolev space  $H_x^s \times H_x^{s-1}$ , where  $s \in [0, 1)$ . However, since this leads to a scaling super-critical problem, all of the above properties can fail. In fact, [16] proved that (1) exhibits norm inflation, which means that arbitrarily small data in  $H^s \times H^{s-1}$  can grow arbitrarily fast. More precisely, we have for all  $\epsilon > 0$  that there exists Schwartz initial data  $(u_0, u_1)$  and a time  $0 < t_\epsilon < \epsilon$  such that  $\|(u_0, u_1)\|_{H^s \times H^{s-1}} < \epsilon$  and  $\|(u(t_\epsilon), \partial_t u(t_\epsilon))\|_{H^s \times H^{s-1}} > \epsilon^{-1}$ . Using finite speed of propagation, one may then also construct solutions whose  $H^s \times H^{s-1}$ -norm blows up instantaneously.

**1.1. The random data Cauchy problem.** Many researchers in dispersive partial differential equations have recently examined whether blow-up behaviour, such as the norm-inflation described above, occurs for generic or only exceptional sets of rough initial data. To quantify this, one is quickly lead to random

initial data. Indeed, one natural form of rough initial data is  $(u_0 + f_0^\omega, u_1 + f_1^\omega)$ , where the functions  $(u_0, u_1) \in \dot{H}^1 \times L^2$  are regular and deterministic, while the functions  $(f_0^\omega, f_1^\omega) \in H^s \times H^{s-1}$  are rough and random. An analogue of Theorem 1.1 in this case would imply the stability of the scattering mechanism under a perturbation by noise.

The literature on random dispersive partial differential equations is vast. We refer the interested reader to the survey [5], and mention the related works [2, 4, 7, 8, 10, 11, 13, 14, 15, 37, 38, 40, 41, 44]. In the following discussion, we focus on the Wiener randomization [3, 37] of a function  $f \in H^s(\mathbb{R}^d)$ . Let  $\varphi \in C^\infty(\mathbb{R}^d)$  be a smooth and symmetric function satisfying  $\varphi|_{[-3/8, 3/8]^d} = 1$ ,  $\varphi|_{\mathbb{R}^d \setminus [-5/8, 5/8]^d} = 0$ , and  $\sum_{k \in \mathbb{Z}^d} \varphi(\xi - k) = 1$  for all  $\xi \in \mathbb{R}^d$ . We then define the associated operator  $P_k$  by

$$\widehat{P_k f}(\xi) := \varphi(\xi - k) \widehat{f}(\xi) .$$

Since the translates  $\{\varphi(\cdot - k)\}_k$  form a partition of unity, we have that

$$(3) \quad f = \sum_{k \in \mathbb{Z}^d} P_k f ,$$

which is called the Wiener decomposition of  $f$ . The Wiener randomization is obtained by randomizing the coefficients in (3). Let  $I \subseteq \mathbb{Z}^d$  be an index set such that  $\mathbb{Z}^d = I \dot{\cup} \{0\} \dot{\cup} (-I)$ . Let  $\{X_k\}_{k \in I \cup \{0\}}$  be a sequence of symmetric, independent, and uniformly sub-gaussian random variables (see Definition 2.1). We set  $X_{-k} := \overline{X_k}$  for all  $k \in I$ , and assume that  $X_0$  is real-valued. Then, the Wiener randomization  $f^W$  is defined as

$$(4) \quad f^W := \sum_{k \in \mathbb{Z}^d} X_k \cdot P_k f .$$

The reason for introducing the set  $I$  is to preserve the real-valuedness of  $f$ . The Wiener randomization  $f^W$  is a random linear combination of functions with unit-scale frequency uncertainty, and therefore resembles a random Fourier series. We then examine the random data Cauchy problem

$$(5) \quad \begin{cases} -\partial_{tt} u + \Delta u = u^3 & (t, x) \in \mathbb{R} \times \mathbb{R}^4 \\ u|_{t=0} = u_0 + f_0^W, \quad \partial_t u|_{t=0} = u_1 + f_1^W \end{cases} .$$

We now seek an almost sure version of Theorem 1.1 for (5). Before we summarize the recent results, let us sketch the overall strategy, which was developed by Pocovnicu in [44]. We let  $F := \cos(t|\nabla|)f_0^W + (\sin(t|\nabla|)/|\nabla|)f_1^W$  be the solution of the linear wave equation with the rough and random initial data. We then define the nonlinear component  $v$  by  $v := u - F$ , and obtain the forced nonlinear wave equation

$$(6) \quad \begin{cases} -\partial_{tt} v + \Delta v = (v + F)^3 & (t, x) \in \mathbb{R} \times \mathbb{R}^4 , \\ v|_{t=0} = u_0, \quad \partial_t v|_{t=0} = u_1 . \end{cases}$$

At the cost of introducing a rough forcing term, we have therefore removed the rough part of the initial data. This transformation is related to the Da Prato-Debussche trick [19]. Due to the smoothing effect of the Duhamel integral, we hope to control the nonlinear component  $v$  in the energy space. The local well-posedness of (6) follows readily from probabilistic Strichartz estimates (cf. [3, 37]) and a contraction mapping argument. Thus our main interest lies in the global well-posedness and the long-time behaviour of the solution. Using the deterministic well-posedness theorem and stability theory, it can be shown (cf.

[26, 44]) that the solution to (6) exists as long as the energy of  $v$  remains bounded. Of course, due to the forcing term in (6), the energy is no longer conserved. In addition, a global bound on the energy of  $v$  implies a global bound on the  $L_t^3 L_x^6$ -norm, and hence also implies scattering. A short calculation shows that

$$(7) \quad \frac{d}{dt} E[v](t) = \int_{\mathbb{R}^4} (v^3 - (v + F)^3) \partial_t v \, dx \approx -3 \int_{\mathbb{R}^4} F v^2 \partial_t v \, dx .$$

In the formula above, we have neglected terms that contain more than a single factor of  $F$ , since they are simpler to estimate. Therefore, the remaining obstacle lies in the control of the right-hand side of (7). With this overall strategy in mind, we summarize the recent literature.

In [44], Pocovnicu proved the almost sure global existence of solutions for all  $s > 0$ . Using a Gronwall-type argument and a probabilistic Strichartz estimate, (7) leads (at top order) to the growth estimate

$$(8) \quad E[v](T) \lesssim E[v](0) \exp(C \|F\|_{L_t^1 L_x^\infty([0,T] \times \mathbb{R}^4)}) \lesssim E[v](0) \exp(C_\omega T^{\frac{1}{2}}) .$$

Since this prevents the finite time blow-up of the energy, this yields an analogue of Theorem 1.1.(i). Similar theorems are also known in dimension five [44], dimension three [43], and for the high-dimensional energy critical nonlinear Schrödinger equation [42].

The bound (8), however, is not sufficient to obtain global control on the energy of  $v$ , and hence does not prove almost sure scattering. Assuming the regularity condition  $s > \frac{1}{2}$  and that the (deterministic) data  $(f_0, f_1)$  is spherically symmetric, Dodson, Lührmann, and Mendelson proved almost sure scattering in [26]. In their argument, the energy increment is estimated by

$$(9) \quad \left| \int_0^T \int_{\mathbb{R}^4} F v^2 \partial_t v \, dx \, dt \right| \lesssim \| |x|^{\frac{1}{2}} F \|_{L_t^2 L_x^\infty([0,T] \times \mathbb{R}^4)} \| |x|^{-\frac{1}{4}} v \|_{L_t^4 L_x^4([0,T] \times \mathbb{R}^4)}^2 \| \partial_t v \|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} .$$

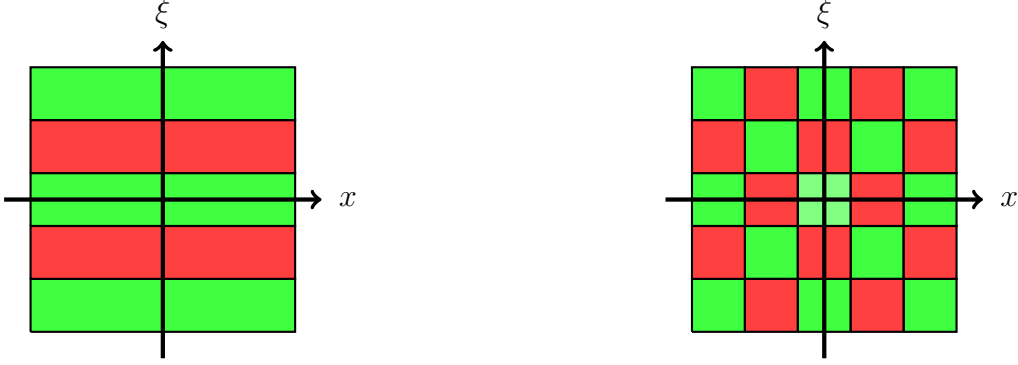
The first factor is controlled using Khintchine's inequality and a square-function estimate, and heavily relies on the spherical symmetry of  $f_0$  and  $f_1$ . The main novelty lies in the control of the second factor, and involves a double bootstrap argument in the energy and a Morawetz term. Under the bootstrap hypothesis, one can then control the second factor in (9) by the square-root of the energy, and this eventually leads to a global bound.

The method of [26] has since been used in several related works. In [25], Dodson, Lührmann, and Mendelson used local energy decay to improve the regularity condition to  $s > 0$ . After replacing the cubes in the Wiener randomization by thin annuli, the author proved almost sure scattering for radial data in dimension three [12]. The main new ingredient is an interaction flux estimate between the linear and nonlinear components of the solution. Finally, the almost sure scattering for the radial energy critical nonlinear Schrödinger equation in four dimensions has been obtained in [25, 32].

**1.2. Main result and ideas.** The remaining open question is concerned with almost sure scattering for non-radial data. In order to state the main result of this paper, we first need to introduce a microlocal randomization. While the Wiener randomization is based on a unit-scale decomposition in frequency space, the microlocal randomization is based on a unit-scale decomposition in phase space (see Figure 1).

**Definition 1.2** (Microlocal randomization).

Let  $\{X_{k,l}\}_{k \in I \cup \{0\}, l \in \mathbb{Z}^d}$  be a sequence of symmetric, independent, and uniformly sub-gaussian random variables. We set  $X_{-k,l} := \overline{X_{k,l}}$  for all  $k \in I$ , and assume that  $X_{0,l}$  is real-valued. For any  $f \in H^s(\mathbb{R}^d)$ , we



We display a partition of the phase space  $\mathbb{R}^d \times \mathbb{R}^d$  into horizontal strips, which forms the basis of the Wiener randomization, and a partition into cubes, which forms the basis of the microlocal randomization. A similar figure has been used in the author's previous work [12, Figure 1].

FIGURE 1. Partions of phase space

define its microlocal randomization  $f^\omega$  by

$$(10) \quad f^\omega(x) := \sum_{k,l \in \mathbb{Z}^d} X_{k,l} P_k(\varphi(\cdot - l)f)(x) .$$

The microlocal randomization is inspired by [39], which used a randomization in physical space.

**Theorem 1.3** (Almost sure scattering for the microlocal randomization).

Let  $(u_0, u_1) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$ , and let  $(f_0, f_1) \in H^s(\mathbb{R}^4) \times H^{s-1}(\mathbb{R}^4)$ , where  $s > \frac{11}{12}$ . Then, there exists a random maximal time interval of existence  $I$  and a solution  $u: I \times \mathbb{R}^4 \rightarrow \mathbb{R}$  of (5) such that

$$u \in W(t)(f_0^\omega, f_1^\omega) + (C_t^0 \dot{H}_x^1(I \times \mathbb{R}^4) \cap L_{t,\text{loc}}^3 L_x^6(I \times \mathbb{R}^4)) \text{ and } \partial_t u \in \partial_t W(t)(f_0^\omega, f_1^\omega) + C_t^0 L_x^2(I \times \mathbb{R}^4) .$$

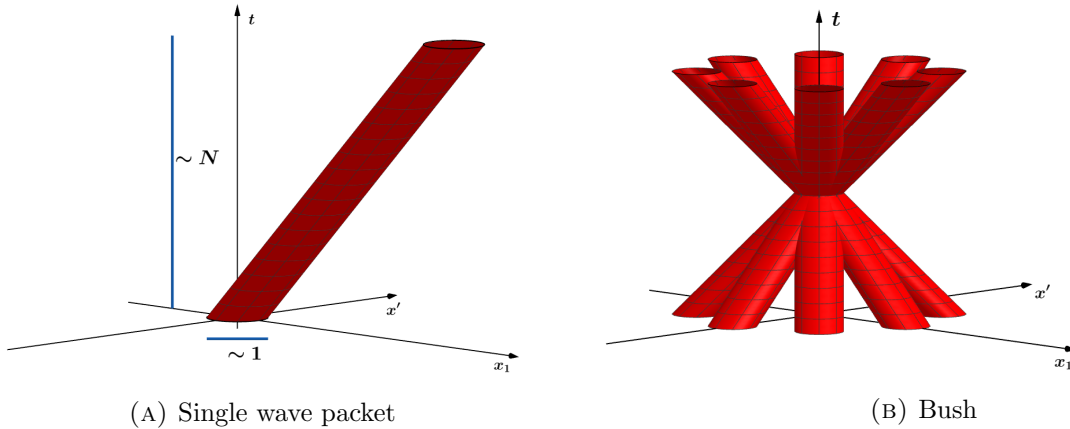
Furthermore, we have that

- (i)  $u$  is almost surely global, i.e.,  $I = \mathbb{R}$ .
- (ii)  $u$  almost surely satisfies the global space-time bound  $\|u\|_{L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)} < \infty$ .
- (iii)  $u$  almost surely scatters to a solution of the linear wave equation. Thus, there exist random scattering states  $(u_0^\pm, u_1^\pm) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  s.t.

$$\lim_{t \rightarrow \pm\infty} \|(u(t) - W(t)(u_0^\pm + f_0^\omega, u_1^\pm + f_1^\omega), \partial_t u(t) - \partial_t W(t)(u_0^\pm + f_0^\omega, u_1^\pm + f_1^\omega))\|_{\dot{H}^1 \times L^2} = 0 .$$

While Theorem 1.3 is only proven for the microlocal randomization, the majority of our argument directly applies to the Wiener randomization.

The main novelty in this paper lies in the application of a wave packet decomposition. To illustrate this idea, fix some  $k \in \mathbb{Z}^d$  with  $\|k\|_\infty \sim N$ , and assume that  $\widehat{f}_k(\xi) = N^{-s} \varphi(\xi - k)$ . Then,  $f_k$  will essentially be unaffected by both the Wiener and microlocal randomizations, and hence forms an important example. From the method of non-stationary phase, it follows for all times  $t \in [0, N]$  that the evolution  $\exp(\pm it|\nabla|)f_k$  is concentrated in the ball  $|x \pm tk/\|k\|_2| \lesssim 1$ , and has amplitude  $\sim N^{-s}$ . In space-time, we



In (a), we display the evolution  $\exp(\pm it|\nabla|)f_k$  on the time-interval  $[0, N]$ . The space-time support can be viewed as a tube of length  $\sim N$  and width  $\sim 1$ . The spatial center travels in a fixed direction at the speed of light, which has been normalized to 1. Furthermore, the amplitude of the evolution is given by  $\sim N^{-s}$ . In (b), we display a so-called bush, which is a collection of wave packets intersecting at a single point.

FIGURE 2. Wave packet heuristic

can therefore view the evolution as a tube, see Figure 2a. For larger times, the dispersion of the evolution becomes significant, and the physical localization deteriorates. The wave packet perspective also explains the effect of the frequency randomization on the evolution. In Figure 2b, we display a bush (cf. [6]), which is a collection of wave packets intersecting at a single point. If all wave packets in the bush have comparable amplitudes and the data is deterministic, one expects that the  $L_t^\infty L_x^\infty$ -norm is proportional to the number of wave packets. For random data, however, the phases of the wave packets are all independent, and the central limit theorem predicts that the  $L_t^\infty L_x^\infty$ -norm should instead be proportional to the square-root of the number of wave packets.

The examples in Figure 2 also illustrates an important heuristic: The natural timescale for the randomized evolution at frequency  $N$  is  $T = N$ . This differs from the natural timescale predicted by the (deterministic) bump-function heuristic, which is  $T = N^{-1}$ . We therefore decompose the positive time-interval as

$$(11) \quad [0, \infty) = \left( \bigcup_{n=0}^{\lfloor N^\theta \rfloor} [nN, (n+1)N) \right) \cup [N^{1+\theta}, \infty),$$

where  $\theta > 0$  is a parameter. Our argument then splits into two separate parts.

On the long-time interval  $[N^{1+\theta}, \infty)$ , we use the additional decay obtained through the physical randomization. The basic idea is that after such a long time, the linear evolution could only be concentrated through constructive interference of a large portion of the initial data, which is highly unlikely due to the physical randomness (see Figure 4). To make this rigorous, we prove an  $L_t^1 L_x^\infty([N^{1+\theta}, \infty) \times \mathbb{R}^4)$ -bound on

$P_N F$ , and this is sufficient to control the energy increment. This part of the proof requires the condition  $s > 1 - \theta/2$ .

The majority of this paper focusses on time intervals such as  $[0, N]$ . This part of the argument does not rely on the physical randomness, and therefore also applies to the Wiener randomization. On this interval, we decompose the evolution into a family of wave packets, see Figure 3. As can be seen from a single wave packet, we cannot (always) control the evolution in  $L_t^1 L_x^\infty$ . Instead, we use the following dichotomy: Either  $F$  consists of only a few wave packets, in which case its support lies on a few light-cones, or it consists of many wave packets, in which case the  $L_t^\infty L_x^\infty$ -norm should be small.

We now present a heuristic and simplified version of the main argument. In order to illustrate the ideas, let us first assume that all wave packets belong to a single frequency  $k \in \mathbb{Z}^d$ . After a dyadic decomposition, we may further assume that all wave packets have amplitudes comparable to  $2^m$ . Using the same notation as in Section 4, we denote the number of wave packets with this amplitude by  $\#\mathcal{A}_m$ . Due to the  $L^2$ -orthogonality of the wave packets, we have that  $2^m (\#\mathcal{A}_m)^{\frac{1}{2}} \lesssim N^{-s}$ .

In the case of only a few wave packets, we control the contribution on each tube separately. We have that

$$\left| \int_0^N F_N v^2 \partial_t v dx dt \right| \lesssim (\#\mathcal{A}_m) N^{\frac{1}{2}} 2^m \left( \sup_{\text{tubes } T} \|v\|_{L_t^4 L_x^4(T)}^2 \right) \|\partial_t v\|_{L_t^\infty L_x^2([0, N] \times \mathbb{R}^4)} \lesssim N^{\frac{1}{2}} 2^m \#\mathcal{A}_m \sup_{t \in [0, N]} E[v](t).$$

The supremum ranges over all tubes of length  $\sim N$ , width  $\sim 1$ , and unit-speed direction inside  $[0, N] \times \mathbb{R}^4$ . Using a flux estimate and a bootstrap argument, we control this supremum by the square-root of the energy.

In the case of many wave packets (with the same direction), we use that their supports are disjoint, and obtain that

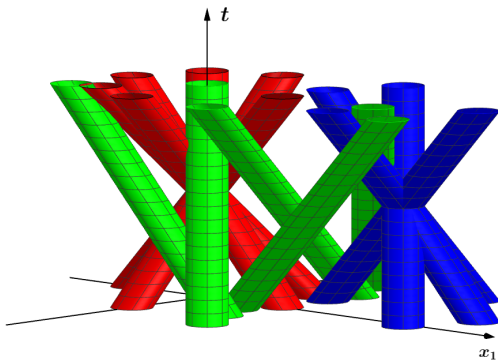
$$\left| \int_0^N F_N v^2 \partial_t v dx dt \right| \lesssim N \|F_N\|_{L_t^\infty L_x^\infty([0, N] \times \mathbb{R}^4)} \|v\|_{L_t^\infty L_x^4([0, N] \times \mathbb{R}^4)}^2 \|\partial_t v\|_{L_t^\infty L_x^2([0, N] \times \mathbb{R}^4)} \lesssim N 2^m \sup_{t \in [0, N]} E[v](t).$$

By combining both estimates, it follows that

$$\left| \int_0^N F_N v^2 \partial_t v dx dt \right| \lesssim \min(N^{\frac{1}{2}} 2^m \#\mathcal{A}_m, N 2^m) \sup_{t \in [0, N]} E[v](t) \leq N^{\frac{3}{4}} 2^m (\#\mathcal{A}_m)^{\frac{1}{2}} \sup_{t \in [0, N]} E[v](t).$$

We insert the bound  $2^m (\#\mathcal{A}_m)^{\frac{1}{2}} \lesssim N^{-s}$ , sum over  $N^\theta$  intervals, and arrive at the condition  $s > 3/4 + \theta$ . In order to match the conditions from the intervals  $[nN, (n+1)N]$  and the long-time interval  $[N^{1+\theta}, \infty)$ , we choose  $\theta = 1/6$ , and obtain the regularity condition  $s > 11/12$ .

In order to remove the restriction to a single frequency, we need to consider both multiple directions and multiple scales. For this, we rely on techniques from the literature on the Keakeya and restriction conjectures. In order to control multiple directions, we use Bourgain's bush argument [6]. The basic idea is to distinguish points which lie in multiple tubes from points which lie only in a few tubes. To this end, we group the wave packets into several bushes and a collection of (almost) non-overlapping wave packets (see Figure 3). We then almost argue as for a single frequency, but also use that each bush lies on the surface of a light-cone, which is crucial for the flux estimate. In order to control multiple scales, we



We illustrate the wave packet decomposition of the linear evolution. We partition the wave packets into three groups: Two separate bushes (red and blue) and a collection of almost non-overlapping wave packets (green).

FIGURE 3. Wave packet decomposition

rely on Wolff's induction on scales strategy [57]. To fix ideas, let us try to bound the energy increment  $E[v](N) - E[v](0)$ . We have already described the estimates for wave packets of length greater than or equal to  $N$ , but the space-time region  $[0, N] \times \mathbb{R}^4$  also contains many shorter wave packets. By induction on scales, we can already close the bootstrap argument at these shorter scales, which greatly reduces the complexity of the proof. We postpone a more detailed discussion to the Sections 4 and 6.

Acknowledgements:

I want to thank my advisor Terence Tao for his invaluable guidance and support. In particular, he proposed the greedy selection algorithm in Section 4. Furthermore, I want to thank Rowan Killip and Monica Visan for several interesting discussions. The figures in this paper have been created using TikZ and GeoGebra.

## 2. NOTATION AND PRELIMINARIES

For the rest of this paper, the positive integer  $d$  denotes the dimension of physical space. In the analysis of the nonlinear evolution, we will eventually specialize to  $d = 4$ . Furthermore, we also fix positive absolute constants  $\delta, \theta$ , and  $\eta$ . The parameter  $\delta$  will be used to deal with the spatial tails of the wave packets and certain summability issues. The parameter  $\theta$  is used in the division of time (see (11)). We will eventually choose  $\theta = 1/6$ , but prefer to keep  $\theta$  as a free parameter until the end of the argument. Finally,  $\eta$  describes the size of the frequency truncated data, see Proposition 4.8.

If  $A, B$  are positive quantities, we write  $A \lesssim B$  if and only if there exist a constant  $C = C(\delta, \theta)$  such that  $A \leq CB$ . Furthermore, most capital letters, such as  $N, M$ , and  $R$ , will denote dyadic numbers greater than or equal to 1.

Finally, we define the Fourier transform  $\hat{f}$  of an integrable function  $f: \mathbb{R}^d \rightarrow \mathbb{C}$  by

$$\hat{f}(\xi) := \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp(-ix\xi) f(x) dx .$$



We now summarize a few basic results from probability theory, harmonic analysis, and dispersive partial differential equations.

**2.1. Probability theory.** We recall a few basic estimates for sub-gaussian random variables. For an accessible introduction, we refer the reader to [54].

**Definition 2.1** (Sub-gaussian random variable).

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space, and let  $X: (\Omega, \mathcal{F}) \rightarrow \mathbb{R}$  be a random variable. We then define the sub-gaussian norm by

$$(12) \quad \|X\|_{\Psi_2} := \sup_{p \geq 1} \frac{(\mathbb{E}[|X|^p])^{\frac{1}{p}}}{\sqrt{p}}$$

We call a random variable  $X$  sub-gaussian if and only if  $\|X\|_{\Psi_2} < \infty$ . Furthermore, we call a family of random variables  $\{X_i\}_{i \in I}$  uniformly sub-gaussian if and only if  $\sup_{i \in I} \|X_i\|_{\Psi_2} < \infty$ .

The relationship to the Gaussian distribution may not be obvious from (12). However, it follows from [54, Proposition 2.52] that (12) implies

$$\mathbb{P}(|X| > x) \leq 2 \exp\left(-c \frac{x^2}{\|X\|_{\Psi_2}^2}\right) \quad \forall x > 0.$$

Many concentration inequalities for the sums of independent sub-gaussian random variables can be found in the literature. In the following, we mainly rely on Khintchine's inequality.

**Lemma 2.2** (Khintchine's inequality, [53, Corollary 5.12] or [54, Proposition 2.6.1 and Exercise 2.6.5]).

Let  $(X_j)_{j=1, \dots, J}$  be a finite sequence of independent sub-gaussian random variables with zero mean. Then, it holds for all deterministic sequences  $(a_j)_{j=1, \dots, J}$ , and all  $p \geq 1$ , that

$$(13) \quad \left(\mathbb{E}\left[\left|\sum_{j=1}^J a_j X_j\right|^p\right]\right)^{\frac{1}{p}} \lesssim \sqrt{p} \left(\max_{j=1, \dots, J} \|X_j\|_{\Psi_2}\right) \left(\sum_{j=1}^J |a_j|^2\right)^{\frac{1}{2}}$$

In particular, the sum  $\sum_{j=1}^J a_j X_j$  is sub-gaussian.

In this paper, Khintchine's inequality will often be combined with Minkowski's integral inequality, which we recall below.

**Lemma 2.3** (Minkowski's integral inequality).

Let  $(X, \mu)$  and  $(Y, \nu)$  be two  $\sigma$ -finite measure spaces, and let  $1 \leq p \leq q \leq \infty$ . Then, we have for all measurable functions  $f: X \times Y \rightarrow \mathbb{R}$  that

$$\| \|f(x, y)\|_{L^p(X)} \|_{L^q(Y)} \leq \| \|f(x, y)\|_{L^q(Y)} \|_{L^p(X)}.$$

The special case  $p = 1$  is the standard Minkowski inequality, and it can be found in most real analysis books (see e.g. [36, Theorem 2.4]). Since Lemma 2.3 is central to many arguments in this paper, we prove the general statement from this special case.

*Proof.* Since  $q/p \geq 1$ , we have that

$$\| \|f(x, y)\|_{L^p(X)} \|_{L^q(Y)} = \| \|f(x, y)^p\|_{L^1(X)}^{\frac{1}{p}} \|_{L^{\frac{q}{p}}(Y)} \leq \| \|f(x, y)^p\|_{L^{\frac{q}{p}}(Y)}^{\frac{1}{p}} \|_{L^1(X)} = \| \|f(x, y)\|_{L^q(Y)} \|_{L^p(X)}.$$

□

We will also need a crude bound on the maximum of dependent sub-gaussian random variables

**Lemma 2.4** (Suprema of dependent sub-gaussian random variables [54, Exercise 2.5.10]).

Assume that  $(X_j)_{j=1,\dots,J}$  are (possibly dependent) sub-gaussian random variables. Then,

$$(14) \quad \mathbb{E} \left[ \max_{j=1,\dots,J} |X_j| \right] \lesssim \sqrt{\log(2+J)} \max_{j=1,\dots,J} \|X_j\|_{\Psi_2} .$$

*Proof.* Let  $1 \leq p < \infty$ . Using Hölder's inequality, we obtain that

$$\mathbb{E} \left[ \max_{j=1,\dots,J} |X_j| \right] \leq \left( \mathbb{E} \left[ \max_{j=1,\dots,J} |X_j|^p \right] \right)^{\frac{1}{p}} \leq \left( \sum_{j=1}^J \mathbb{E} [|X_j|^p] \right)^{\frac{1}{p}} \leq J^{\frac{1}{p}} \sqrt[p]{p} \max_{j=1,\dots,J} \|X_j\|_{\Psi_2} .$$

After choosing  $p = \log(2+J)$ , we arrive at (14).  $\square$

**2.2. Harmonic analysis.** Let  $N \in 2^{\mathbb{N}_0}$  and  $k \in \mathbb{Z}^d$ . As in the introduction, we let  $\varphi \in C_c^\infty(\mathbb{R}^d)$  be a smooth and symmetric function satisfying  $\varphi|_{[-\frac{3}{8}, \frac{3}{8}]^d} = 1$ ,  $\varphi|_{\mathbb{R}^d \setminus [-\frac{5}{8}, \frac{5}{8}]^d} = 0$ , and  $\sum_{k \in \mathbb{Z}^d} \varphi(\cdot - k) = 1$ . We also define  $\psi(\xi) = \varphi(\xi) - \varphi(2\xi)$ . Then, the re-centered Littlewood-Paley operators are defined as

$$\widehat{P_{N;k}f}(\xi) = \begin{cases} \psi\left(\frac{\xi-k}{N}\right) \widehat{f}(\xi) & \text{if } N > 1 \\ \varphi(\xi - k) \widehat{f}(\xi) & \text{if } N = 1 \end{cases} .$$

With this choice of  $\psi$ , it holds that  $P_{N;k}P_{M;k} = 0$  if  $M \geq 4N$  or  $N \geq 4M$ . To simplify the notation, we also set  $P_N := P_{N;0}$  and  $P_k := P_{1;k}$ . Furthermore, we define the fattened Littlewood-Paley operators

$$(15) \quad \widetilde{P}_{N;k} := \sum_{2^{-10}N \leq M \leq 2^{10}N} P_{M;k} .$$

**Lemma 2.5** (Bernstein's inequalities).

Let  $N \in 2^{\mathbb{N}_0}$ ,  $k \in \mathbb{Z}^d$ , and  $1 \leq p \leq q \leq \infty$ . Then, we for all  $f \in L_x^p(\mathbb{R}^d)$  that

$$(16) \quad \|P_{N;k}f\|_{L_x^q(\mathbb{R}^d)} \lesssim N^{d(\frac{1}{p}-\frac{1}{q})} \|P_{N;k}f\|_{L_x^p(\mathbb{R}^d)} ,$$

$$(17) \quad \|\nabla|P_{N;0}f\|_{L_x^p(\mathbb{R}^d)} \lesssim N \|P_{N;0}f\|_{L_x^p(\mathbb{R}^d)} .$$

We emphasize that the constant in (16) is independent of  $k \in \mathbb{Z}^d$ , since the phase  $\exp(ikx)$  does not affect the  $L_x^p$ -norms.

We also record the following standard consequence of Bernstein's inequality and the uncertainty principle.

**Lemma 2.6.**

Let  $N \geq 1$ , let  $a, b \in \mathbb{R}$ , and assume that  $b \geq a + 1/N$ . Then, we have for all  $f \in L_x^2(\mathbb{R}^d)$ , all  $1 \leq q < \infty$ , and all  $1 \leq p \leq \infty$ , that

$$\|\exp(\pm it|\nabla|)P_N f\|_{L_t^\infty L_x^p([a,b] \times \mathbb{R}^d)} \lesssim N^{\frac{1}{q}} \|\exp(\pm it|\nabla|)P_N f\|_{L_t^q L_x^p([a,b] \times \mathbb{R}^d)} .$$

The argument is essentially taken from [12].

*Proof.* Pick  $t_0 \in [a, b]$ , and let  $I$  be any interval such that  $t_0 \in I \subseteq [a, b]$ . For all  $t \in I$ , it holds that

$$\exp(\pm it_0 |\nabla|) P_N f^\omega = \exp(\pm it |\nabla|) P_N f^\omega \pm i \int_t^{t_0} |\nabla| \exp(\pm it' |\nabla|) P_N f^\omega dt'$$

From Bernstein's inequality, we obtain that

$$\begin{aligned} \|\exp(\pm it_0 |\nabla|) P_N f^\omega\|_{L_x^p(\mathbb{R}^d)} &\leq \|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_x^p(\mathbb{R}^d)} + |I|^{1-\frac{1}{q}} \|\nabla \exp(\pm it |\nabla|) P_N f^\omega\|_{L_t^q L_x^p(I \times \mathbb{R}^d)} \\ &\lesssim \|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_x^p(\mathbb{R}^d)} + N |I|^{1-\frac{1}{q}} \|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_t^q L_x^p(I \times \mathbb{R}^d)} \end{aligned}$$

Taking the  $q$ -th power and averaging over all  $t \in I$ , we obtain that

$$\|\exp(\pm it_0 |\nabla|) P_N f^\omega\|_{L_x^p(\mathbb{R}^d)} \lesssim \left( |I|^{-\frac{1}{q}} + N |I|^{1-\frac{1}{q}} \right) \|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_t^q L_x^p(I \times \mathbb{R}^d)} .$$

By choosing  $|I| = N^{-1}$ , and taking the supremum over all  $t_0 \in [a, b]$ , it follows that

$$\|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_t^\infty L_x^p([a, b] \times \mathbb{R}^d)} \lesssim N^{\frac{1}{q}} \|\exp(\pm it |\nabla|) P_N f^\omega\|_{L_t^q L_x^p([a, b] \times \mathbb{R}^d)}$$

□

The following estimate, which appeared in the almost sure scattering problem for the nonlinear Schrödinger equation [32], is useful in combination with Khintchine's inequality.

**Lemma 2.7** (Square function estimate [32, Lemma 2.8]).

Let  $f \in L_x^2(\mathbb{R}^d)$  and let  $\varphi$  be as above. Then, it holds that

$$(18) \quad \sum_{k \in \mathbb{Z}^d} |P_k f(x)|^2 \lesssim (|\check{\varphi}| * |f|^2)(x) .$$

In addition to the dyadic decomposition in frequency, we also need a dyadic decomposition in physical space. To avoid confusion, we denote the cut-off function in physical space by  $\chi$ . More precisely, we set  $\chi_1(x) := \varphi(x)$  and  $\chi_L(x) := \psi(x/L)$ , where  $L \geq 2$ .

**Lemma 2.8** (Mismatch estimates).

Let  $1 \leq p \leq \infty$  and  $M, N, L \geq 1$ . We further assume the mismatch conditions  $\max(M/N, N/M) \geq 8$  and  $L \geq 8$ . Then, it holds for all absolute constants  $D > 0$  that

$$(19) \quad \|\chi_1 P_M \chi_L\|_{L_x^p \rightarrow L_x^p} \lesssim_D (ML)^{-D} ,$$

$$(20) \quad \|P_N \chi_1 P_M\|_{L_x^p \rightarrow L_x^p} \lesssim_D (NM)^{-D} .$$

*Proof.* The inequality (19) can be found in [25, Lemma 5.10]. An inequality similar to (20) can be found in [25, Lemma 5.11], and we present a different argument.

Using duality and  $(P_N \chi_1 P_M)^* = P_M \chi_1 P_N$ , we can assume that  $N \geq M$ . From the mismatch condition, it then follows that  $N \geq 8M$ . Thus, we obtain for all  $f \in L_x^p(\mathbb{R}^d)$  that

$$\|P_N(\chi_1 P_M f)\|_{L_x^p} = \|P_N((P_{\geq N/8} \chi_1) P_M f)\|_{L_x^p} \lesssim \|P_{\geq N/8} \chi_1\|_{L_x^\infty} \|P_M f\|_{L_x^p} \lesssim N^{-2D} \|f\|_{L_x^p} .$$

□

The following auxiliary lemma will be helpful in the proof of probabilistic Strichartz estimates.

**Lemma 2.9** ( $\ell_{k,l}^2$ -estimate).

Let  $s \in \mathbb{R}$  and let  $f \in H_x^s(\mathbb{R}^d)$ . For any  $2 \leq p \leq \infty$ , we have that

$$(21) \quad \|P_k(\varphi_l f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}(\mathbb{Z}^d \times \{k \in \mathbb{Z}^d : \|k\|_\infty \in (N/2, N]\} \times \mathbb{R}^d)} \lesssim \|\tilde{P}_N f\|_{L_x^2(\mathbb{R}^d)} + N^{-s-10d} \|f\|_{H_x^s(\mathbb{R}^d)}$$

**Remark 2.10.**

The error term  $N^{-s-10d} \|f\|_{H_x^s(\mathbb{R}^d)}$  is a result of the non-compact support of  $\widehat{\varphi}_l$ , but may essentially be ignored. On a heuristic level, each  $P_k(\varphi_l f)$  is supported on a spatial region of volume  $\sim 1$ , and thus (21) should follow from Hölder's inequality. To make this argument rigorous, we use the square-function estimate and the mismatch estimates above.

*Proof.* Let  $\tilde{P}_N$  be the fattened Littlewood-Paley operator as in (15). We write  $M \not\asymp N$  if either  $M < 2^{-10}N$  or  $M > 2^{10}N$ . In the following, we implicitly assume that  $\|k\|_\infty \in (N/2, N]$ . We then estimate

$$(22) \quad \|P_k(\varphi_l f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}} \leq \|P_k(\varphi_l \tilde{P}_N f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}} + \sum_{M \not\asymp N} \|P_k(\varphi_l P_M f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}}.$$

We begin by controlling the first summand in (22). Using Minkowski's integral inequality and the square-function estimate (Lemma 2.7), we obtain that

$$(23) \quad \begin{aligned} & \|P_k(\varphi_l \tilde{P}_N f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}} = \|\varphi_{l'} P_k(\varphi_l \tilde{P}_N f)\|_{\ell_k^2 \ell_k^2 \ell_{l'}^1 L_x^{p'}} \lesssim \|\varphi_{l'} P_k(\varphi_l \tilde{P}_N f)\|_{\ell_k^2 \ell_k^2 \ell_{l'}^1 L_x^2} \lesssim \|\varphi_{l'} P_k(\varphi_l \tilde{P}_N f)\|_{\ell_{l'}^1 \ell_k^1 L_x^2 \ell_k^2} \\ & = \|\varphi_{l'} \left( |\check{\varphi}| * |\varphi_l \tilde{P}_N f|^2 \right)^{\frac{1}{2}}\|_{\ell_{l'}^1 \ell_{l'}^1 L_x^2}. \end{aligned}$$

Using simple support considerations, we have that

$$\begin{aligned} & \|\varphi_{l'} \left( |\check{\varphi}| * |\varphi_l \tilde{P}_N f|^2 \right)^{\frac{1}{2}}\|_{L_x^2}^2 \leq \|\varphi_{l'} \left( |\check{\varphi}| * |\varphi_l \tilde{P}_N f|^2 \right)\|_{L_x^1} \lesssim \langle l' - l \rangle^{-10d} \|(\varphi_l \tilde{P}_N f)^2\|_{L_x^1} \\ & \lesssim \langle l' - l \rangle^{-10d} \|\varphi_l \tilde{P}_N f\|_{L_x^2}^2. \end{aligned}$$

Inserting this back into (23), we obtain that

$$\|P_k(\varphi_l \tilde{P}_N f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}} \lesssim \langle l' - l \rangle^{-5d} \|\varphi_l \tilde{P}_N f\|_{\ell_k^2 \ell_k^1 L_x^2} \lesssim \|\tilde{P}_N f\|_{L_x^2}.$$

Thus, this yields the first term in (21). We now control the second summand in (22). First, note that  $P_k = \sum_{2^{-5}N \leq N' \leq 2^5 N} P_{N'} P_k$ . Since there exist only  $\sim N^d$  frequencies of magnitude  $\sim N$ , we have that

$$\|P_k(\varphi_l P_M f)\|_{\ell_k^2 \ell_k^2 L_x^{p'}} \lesssim N^{\frac{d}{2}} \sum_{2^{-5}N \leq N' \leq 2^5 N} \|P_{N'}(\varphi_l P_M f)\|_{\ell_{l'}^2 L_x^{p'}}.$$

It now suffices to prove for all  $g \in S(\mathbb{R}^d)$ , all  $M \not\asymp N$ , and all absolute constants  $D > 0$  that

$$(24) \quad \|P_{N'}(\varphi_l P_M g)\|_{L_x^{p'}} \lesssim_D (NM)^{-D} \|\langle x - l \rangle^{-D} g\|_{L_x^2}.$$

Using spatial translation invariance, we may set  $l = 0$ . Let  $\{\chi_L\}_{L \geq 1}$  denote the dyadic decomposition in physical space. Using the mismatch estimates (Lemma 2.8), we obtain

$$\begin{aligned} \|P_{N'}(\varphi_0 P_M g)\|_{L_x^{p'}} &\leq \sum_{L \geq 1} \|P_{N'}(\varphi_0 P_M \chi_L g)\|_{L_x^{p'}} \leq \sum_{L \geq 1} \|P_{N'} \varphi_0 P_M \chi_L\|_{L_x^{p'} \rightarrow L_x^{p'}} \|\tilde{\chi}_L g\|_{L_x^{p'}} \\ &\lesssim (NM)^{-D} \sum_{L \geq 1} L^{-2D} \|\tilde{\chi}_L g\|_{L_x^{p'}} \lesssim (NM)^{-D} \langle x \rangle^{-D} g \|_{L_x^2} . \end{aligned}$$

□

As a direction consequence of (2.9), we also obtain the following estimate on the  $H^s$ -norm of the microlocal randomization.

**Lemma 2.11** ( $H_x^s$ -norm of  $f^\omega$ ).

Let  $f \in H_x^s(\mathbb{R}^d)$  and let  $f^\omega$  be its microlocal randomization. We further set

$$(25) \quad f_1^\omega := \sum_{l \in \mathbb{Z}^d} X_{0,l} P_0(\varphi_l f) \quad \text{and} \quad f_N^\omega := \sum_{\substack{k,l \in \mathbb{Z}^d \\ \|k\|_\infty \in (N/2, N]}} X_{k,l} P_k(\varphi_l f) , \quad \text{where } N \geq 2 .$$

Then, we have for all  $2 \leq r < \infty$  that

$$(26) \quad \|f^\omega\|_{L_\omega H_x^s} \simeq \|N^s f_N^\omega\|_{L_\omega \ell_N^2 L_x^2} \lesssim \sqrt{r} \|f\|_{H_x^s} .$$

*Proof.* The first equivalence in (26) is a direct consequence of the definition of the  $H_x^s$ -norm. Now, we prove the bound in (26). From Minkowski's integral inequality, Khintchine's inequality, and Lemma 2.9, we have for all  $N \geq 2$  that

$$\begin{aligned} \|N^s f_N^\omega\|_{L_\omega L_x^2} &\leq \|N^s \sum_{\|k\|_\infty \in (N/2, N]} X_{k,l} P_k(\varphi_l f)\|_{L_x^2 L_\omega} \\ &\leq \sqrt{r} N^s \|P_k(\varphi_l f)\|_{L_x^2 \ell_{k,l}^2(\|k\|_\infty \in (N/2, N])} \\ &\lesssim \sqrt{r} \left( N^s \|\tilde{P}_N f\|_{L_x^2(\mathbb{R}^d)} + N^{-10d} \|f\|_{H_x^s(\mathbb{R}^d)} \right) \end{aligned}$$

The same argument also applies to  $N = 1$ . After taking the  $\ell_N^2$ -norm, this completes the proof. □

**2.3. Strichartz estimates.** The individual blocks in the microlocal randomization or the Wiener randomization have frequency support inside a unit-sized cube (at a large distance from the origin). Since this rules out the Knapp example, one expects a refined dispersive estimate. The following lemma is due to Klainerman and Tataru [35], and it has first been used in the probabilistic context by [26].

**Lemma 2.12** (Refined dispersive estimate by Klainerman-Tataru [35]).

Let  $f \in L^1(\mathbb{R}^d)$ , let  $k \in \mathbb{Z}^d$  satisfy  $\|k\|_\infty \in (N/2, N]$ , and let  $M \leq N$ . Then it holds for all  $t \in \mathbb{R}$  and  $2 \leq p \leq \infty$  that

$$(27) \quad \|\exp(\pm it|\nabla|) P_{M;k} f\|_{L_x^p(\mathbb{R}^d)} \lesssim \frac{M^{d(1-\frac{2}{p})}}{(1 + \frac{M^2}{N}|t|)^{(d-1)(\frac{1}{2}-\frac{1}{p})}} \|f\|_{L_x^{p'}(\mathbb{R}^d)} .$$

As stated, the inequality (27) essentially follows from [35]. For the sake of completeness, we present the modification below.

*Proof.* By interpolation against the energy estimate  $\|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_x^2(\mathbb{R}^d)} \leq \|f\|_{L_x^2(\mathbb{R}^d)}$ , it suffices to prove (27) for  $p = \infty$ . The inequality [35, (A.66)], where  $\mu = M/N$ , and a scaling argument yield

$$(28) \quad \|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_x^\infty(\mathbb{R}^d)} \lesssim \frac{MN^{d-1}}{(1+N|t|)^{\frac{d-1}{2}}} \|f\|_{L_x^1(\mathbb{R}^d)}.$$

We now distinguish two cases. If  $|t| \lesssim N/M^2$ , then Bernstein's inequality (Lemma 2.5) yields that

$$\|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_x^\infty(\mathbb{R}^d)} \lesssim M^{\frac{d}{2}} \|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_x^2(\mathbb{R}^d)} = M^{\frac{d}{2}} \|P_{M;k}f\|_{L_x^2(\mathbb{R}^d)} \lesssim M^d \|P_{M;k}f\|_{L_x^1(\mathbb{R}^d)}.$$

If  $|t| \gtrsim N/M^2$ , then (28) yields that

$$\|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_x^\infty(\mathbb{R}^d)} \lesssim \frac{MN^{d-1}}{(N|t|)^{\frac{d-1}{2}}} \|f\|_{L_x^1(\mathbb{R}^d)} = \frac{M^d}{(\frac{M^2}{N}t)^{\frac{d-1}{2}}} \|f\|_{L_x^1(\mathbb{R}^d)} \lesssim \frac{M^d}{(1+\frac{M^2}{N}t)^{\frac{d-1}{2}}} \|f\|_{L_x^1(\mathbb{R}^d)}.$$

□

In this paper, we are mainly concerned with the case  $M = 1$ . Then, (27) describes the linear evolution on short time intervals more accurately than (28). As a corollary of the refined dispersive estimate, we obtain the following refined Strichartz estimate.

Let  $2 \leq q, p \leq \infty$ . We call the pair  $(q, p)$  wave-admissible if

$$\frac{1}{q} + \frac{d-1}{2p} \leq \frac{d-1}{4} \quad \text{and} \quad (q, p, d) \neq (2, \infty, 3).$$

**Corollary 2.13** (Refined Strichartz estimates [35]).

Let  $f \in L_x^2(\mathbb{R}^d)$ , let  $k \in \mathbb{Z}^d$  satisfy  $\|k\|_\infty \in (N/2, N]$ , and let  $M \leq N$ . Then, we have for all wave-admissible pairs  $(q, p)$  that

$$(29) \quad \|\exp(\pm it|\nabla|)P_{M;k}f\|_{L_t^q L_x^p(\mathbb{R} \times \mathbb{R}^d)} \lesssim M^{\frac{d}{2} - \frac{2}{q} - \frac{d}{p}} N^{\frac{1}{q}} \|P_{M;k}f\|_{L_x^2(\mathbb{R}^d)}$$

The derivation of the refined Strichartz estimate from Lemma 2.12 follows from a standard  $TT^*$ -argument, and we therefore omit the proof. For the endpoint  $(2, 2(d-1)/(d-3))$ , we also refer to [31]. Let us emphasize two special cases: If  $M = N$ , we obtain the usual scaling factor  $N^{\frac{d}{2} - \frac{1}{q} - \frac{d}{p}}$ , and if  $M = 1$ , we obtain the factor  $N^{\frac{1}{q}}$ , which does not depend on  $p$ .

### 3. PROBABILISTIC STRICHARTZ ESTIMATES

In this section, we derive probabilistic Strichartz estimates (cf. [3, 26, 37]) and a probabilistic long-time decay estimate (cf. [39]). To keep the exposition self-contained, we include the (short) proofs. Recall from (25) that

$$f_1^\omega := \sum_{l \in \mathbb{Z}^d} X_{0,l} P_0(\varphi_l f) \quad \text{and} \quad f_N^\omega := \sum_{\substack{k, l \in \mathbb{Z}^d \\ \|k\|_\infty \in (N/2, N]}} X_{k,l} P_k(\varphi_l f), \quad \text{where } N \geq 2.$$

**Lemma 3.1** (Probabilistic Strichartz estimate).

Let  $f \in H_x^s(\mathbb{R}^d)$  and let  $f^\omega$  be its microlocal randomization. Then, it holds for all  $N \geq 1$ , all wave-admissible exponent pairs  $(q, p)$ , and all  $1 \leq r < \infty$  that

$$(30) \quad \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^r L_t^q L_x^p(\Omega \times \mathbb{R} \times \mathbb{R}^d)} \lesssim \sqrt{r} N^{\frac{1}{q}-s+} \|f\|_{H_x^s(\mathbb{R}^d)} .$$

This estimate has previously appeared for the Wiener randomization in [26].

*Proof.* In the following, we implicitly assume that  $k \in \mathbb{Z}^d$  always satisfies  $\|k\|_\infty \in (N/2, N]$ . First, we assume that  $2 \leq p, q < \infty$ , and that  $\max(p, q) \leq r < \infty$ . Using Minkowski's integral inequality (Lemma 2.3), Khintchine's inequality (Lemma 2.2), and the refined dispersive estimate (Lemma 2.12), we have that

$$\begin{aligned} \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^r L_t^q L_x^p} &\leq \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_t^q L_x^p L_\omega^r} \lesssim \sqrt{r} \|\exp(\pm it|\nabla|)P_k(\varphi_l f)\|_{L_t^q L_x^p \ell_{k,l}^2} \\ &\leq \sqrt{r} \|\exp(\pm it|\nabla|)P_k(\varphi_l f)\|_{\ell_{k,l}^2 L_t^q L_x^p} \lesssim \sqrt{r} N^{\frac{1}{q}} \|P_k(\varphi_l f)\|_{\ell_{k,l}^2 L_x^2} \lesssim \sqrt{r} N^{\frac{1}{q}-s} \|f\|_{H_x^s} . \end{aligned}$$

In the last inequality, we have also used Lemma 2.9. The estimate for  $1 \leq r \leq \max(p, q)$  then follows from Hölder's inequality. Thus, it remains to treat the cases  $q = \infty$  and/or  $p = \infty$ . This is a known technical issue, see [12, Remark 3.8] for a discussion. Both cases can be reduced to the previous estimate by using Lemma 2.6 and Bernstein's inequality.  $\square$

**Lemma 3.2** (Probabilistic long-time decay).

Let  $f \in L_x^2(\mathbb{R}^d)$  and let  $f^\omega$  be its microlocal randomization. Furthermore, let  $1 \leq q < \infty$  and  $2 \leq p \leq \infty$  be such that

$$(31) \quad \frac{1}{q} + \frac{d-1}{p} < \frac{d-1}{2} .$$

Then, we have for all  $1 \leq r < \infty$  that

$$(32) \quad \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^r L_t^q L_x^p(\Omega \times [T, \infty) \times \mathbb{R}^d)} \lesssim_{q,p} \sqrt{r} N^{\frac{1}{q}-s+} \left(1 + \frac{T}{N}\right)^{\frac{1}{q} + \frac{d-1}{p} - \frac{d-1}{2}} \|f\|_{H_x^s(\mathbb{R}^d)} .$$

Lemma 3.2 has previously been used for a physical space randomization in [39, Proposition 3.1]. In contrast to the standard Strichartz estimates, which are time-translation invariant, (32) provides a quantitative decay rate. The motivation behind this estimate is illustrated in Figure 4. In this paper, we only require the following special case.

**Corollary 3.3.**

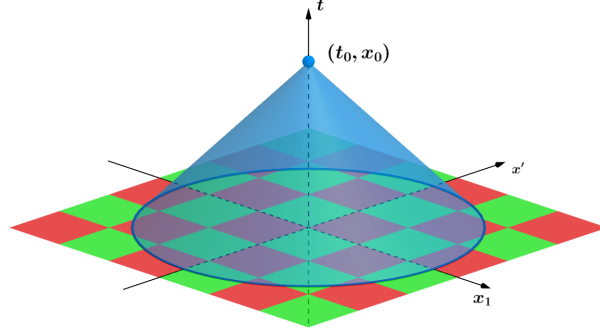
Let  $f \in L_x^2(\mathbb{R}^4)$  and let  $f^\omega$  be its microlocal randomization. Then, we have for all  $\theta > 0$  that

$$(33) \quad \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^r L_t^1 L_x^\infty(\Omega \times [N^{1+\theta}, \infty) \times \mathbb{R}^4)} \lesssim \sqrt{r} N^{1-\frac{\theta}{2}+} \|f_N\|_{L_x^2(\mathbb{R}^4)} .$$

**Remark 3.4.**

Due to (31), the  $L_t^1 L_x^\infty$ -estimate fails logarithmically in three dimensions.

*Proof of Lemma 3.2.* We essentially follow the argument in [39]. Let us first assume that  $2 \leq q, p < \infty$ . We further assume that  $r \geq \max(q, p)$ , the corresponding estimate for  $1 \leq r < \max(q, p)$  then follows from



This figure illustrates the effect of the physical randomization on the linear evolution. At the point  $(t_0, x_0)$ , the linear evolution depends on the initial data in a large region of space. Due to the physical randomization, the initial data in different spatial regions cannot constructively interfere, and hence we expect an improved decay.

FIGURE 4. Effect of physical randomization

Hölder's inequality. Using Minkowski's integral inequality (Lemma 2.3), Khintchine's inequality (Lemma 2.2), and the refined dispersive estimate (Lemma 2.12), we have that

$$\begin{aligned}
& \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_t^\infty L_t^q L_x^p(\Omega \times [T, \infty) \times \mathbb{R}^d)} \\
& \leq \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_t^q L_x^p L_\omega^\infty([T, \infty) \times \mathbb{R}^d \times \Omega)} \\
& \lesssim \sqrt{r} \|\exp(\pm it|\nabla|)P_k(\varphi_l f)\|_{L_t^q L_x^p \ell_{k,l}^2([T, \infty) \times \mathbb{R}^d \times \mathbb{Z}^{d+d})} \\
& \lesssim \sqrt{r} \|\exp(\pm it|\nabla|)P_k(\varphi_l f)\|_{\ell_{k,l}^2 L_t^q L_x^p(\mathbb{Z}^{d+d} \times [T, \infty) \times \mathbb{R}^d)} \\
& \lesssim \sqrt{r} \left\| \left(1 + \frac{|t|}{N}\right)^{-(d-1)(\frac{1}{2} - \frac{1}{p})} \|P_k(\varphi_l f)\|_{L_x^{p'}} \right\|_{\ell_{k,l}^2 L_t^q(\mathbb{Z}^{d+d} \times [T, \infty))} \\
& \lesssim \sqrt{r} \left\| \left(1 + \frac{|t|}{N}\right)^{-(d-1)(\frac{1}{2} - \frac{1}{p})} \|P_k(\varphi_l f)\|_{\ell_{k,l}^2 L_x^{p'}(\mathbb{Z}^{d+d} \times \mathbb{R}^d)} \right\|_{L_t^q([T, \infty))}.
\end{aligned}$$

Using condition (31), we obtain

$$\left\| \left(1 + \frac{|t|}{N}\right)^{-(d-1)(\frac{1}{2} - \frac{1}{p})} \right\|_{L_t^q([T, \infty))} \lesssim N^{\frac{1}{q}} \left(1 + \frac{T}{N}\right)^{\frac{1}{q} + \frac{d-1}{p} - \frac{d-1}{2}}$$

Finally, from Lemma 2.9 we have that

$$\|P_k(\varphi_l f)\|_{\ell_{k,l}^2 L_x^{p'}(\mathbb{Z}^{d+d} \times \mathbb{R}^d)} \lesssim N^{-s} \|f\|_{H_x^s}.$$

This finishes the proof in the case  $2 \leq q, p < \infty$ . Using Bernstein's inequality, we can reduce the case  $p = \infty$  to  $p < \infty$ . Thus, it remains to treat the range  $1 \leq q < 2$ . Using a dyadic decomposition in time,



we have for all  $T \geq N$  that

$$\begin{aligned}
& \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^2 L_t^q L_x^p(\Omega \times [T, \infty) \times \mathbb{R}^d)} \\
& \leq \sum_{m=0}^{\infty} \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^2 L_t^q L_x^p(\Omega \times [2^m T, 2^{m+1} T) \times \mathbb{R}^d)} \\
& \leq \sum_{m=0}^{\infty} (2^m T)^{\frac{1}{q}-\frac{1}{2}} \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^2 L_t^2 L_x^p(\Omega \times [2^m T, 2^{m+1} T) \times \mathbb{R}^d)} \\
& \lesssim \sum_{m=0}^{\infty} (2^m T)^{\frac{1}{q}-\frac{1}{2}} N^{\frac{1}{2}-s} \left(\frac{2^m T}{N}\right)^{\frac{1}{2}+\frac{d-1}{p}-\frac{d-1}{2}} \|f\|_{H_x^s} \\
& \lesssim N^{\frac{1}{q}-s} \left(\frac{T}{N}\right)^{\frac{1}{q}+\frac{d-1}{p}-\frac{d-1}{2}} \|f\|_{H_x^s}.
\end{aligned}$$

In the second last line, we used condition (31). For  $T \leq N$ , we also have that

$$\|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^2 L_t^q L_x^p(\Omega \times [T, N) \times \mathbb{R}^d)} \lesssim N^{\frac{1}{q}-\frac{1}{2}} \|\exp(\pm it|\nabla|)f_N^\omega\|_{L_\omega^2 L_t^2 L_x^p(\Omega \times [0, N) \times \mathbb{R}^d)} \lesssim N^{\frac{1}{q}-s} \|f\|_{H_x^s}.$$

□

**Definition 3.5** (Auxiliary norm).

Let  $0 \leq s < 1$ , let  $(f_0, f_1) \in H^s(\mathbb{R}^4) \times H^{s-1}(\mathbb{R}^4)$ , and let  $N_0 \geq 1$ . We then define

$$\begin{aligned}
\|(f_0, f_1)\|_{Z(N_0)} & := \sum_{N \geq N_0} N^{s+\frac{\theta}{2}-1-\delta} \left\| \cos(t|\nabla|)f_{0,N} + \frac{\sin(t|\nabla|)}{|\nabla|} f_{1,N} \right\|_{L_t^1 L_x^\infty([N^{1+\theta}, \infty) \times \mathbb{R}^4)} \\
& + \sum_{N \geq N_0} N^{s-\delta} \left\| \cos(t|\nabla|)f_{0,N} + \frac{\sin(t|\nabla|)}{|\nabla|} f_{1,N} \right\|_{L_t^\infty L_x^\infty([0, \infty) \times \mathbb{R}^4)}.
\end{aligned}$$

From Proposition 3.1 and Corollary 3.3, it follows that

$$\|(f_0^\omega, f_1^\omega)\|_{L_\omega^2 Z(1)} \lesssim \sqrt{r} \|(f_0, f_1)\|_{H_x^s \times H_x^{s-1}}.$$

#### 4. WAVE PACKET DECOMPOSITION

In this section, we use a wave packet decomposition to better understand the (random) linear evolution. This part of the argument does not rely on the additional randomization in physical space. We therefore phrase all results in a way that applies to both the microlocal and the Wiener randomization, and hope that this facilitates future applications. With this in mind, we now rewrite the microlocal randomization in a form that resembles the Wiener randomization.

Let the random variables  $\{X_{k,l}\}_{k,l \in \mathbb{Z}^d}$  be as in Definition 1.2, let  $\{\epsilon_k\}_{k \in I \cup \{0\}}$  be a family of independent random signs, and set  $\epsilon_{-k} = \epsilon_k$  for all  $k \in I$ . We can then define  $Y_{k,l} := \epsilon_k X_{k,l}$ . For a sequence of

multi-indices  $k, l_1, \dots, l_J \in \mathbb{Z}^d$  and any sequence of Borel-measurable sets  $A_1, \dots, A_J \subseteq \mathbb{R}$ , we have that

$$\begin{aligned} \mathbb{P}(\epsilon_k = 1, Y_{k,l_1} \in A_1, \dots, Y_{k,l_J} \in A_J) &= \mathbb{P}(\epsilon_k = 1, X_{k,l_1} \in A_1, \dots, X_{k,l_J} \in A_J) \\ &= \mathbb{P}(\epsilon_k = 1) \prod_{j=1}^J \mathbb{P}(X_{k,l_j} \in A_j) = \mathbb{P}(\epsilon_k = 1) \prod_{j=1}^J \mathbb{P}(Y_{k,l_j} \in A_j) . \end{aligned}$$

In the last equality, we have used that the random variables  $X_{k,l}$  are symmetric. Therefore, for a fixed  $k \in \mathbb{Z}^d$ , the family  $\{\epsilon_k\} \cup \{Y_{k,l}\}_{l \in \mathbb{Z}^d}$  is independent. From this, it then easily follows that the whole family  $\{\epsilon_k\}_{k \in I \cup \{0\}} \cup \{Y_{k,l}\}_{k \in I \cup \{0\}, l \in \mathbb{Z}^d}$  is independent. We then rewrite the microlocal randomization as

$$(34) \quad f^\omega = \sum_{k,l \in \mathbb{Z}^d} X_{k,l} P_k(\varphi_l f) = \sum_{k \in \mathbb{Z}^d} \epsilon_k P_k \left( \sum_{l \in \mathbb{Z}^d} Y_{k,l} \varphi_l f \right) = \sum_{k \in \mathbb{Z}^d} \epsilon_k f_k, \quad \text{where } f_k := P_k \left( \sum_{l \in \mathbb{Z}^d} Y_{k,l} \varphi_l f \right) .$$

Due to the independence properties discussed above, we can regard the functions  $\{f_k\}$  as deterministic by conditioning on the random variables  $\{Y_{k,l}\}_{k,l}$ , and only utilize the randomness through the random signs  $\{\epsilon_k\}_k$ . Note that (34) closely resembles the Wiener randomization.

To motivate the wave packet decomposition below, we now rewrite the linear evolution with initial data  $(f_0^\omega, f_1^\omega)$ . Using the notation from (34), we first introduce the half-wave operators by writing

$$\begin{aligned} &\cos(t|\nabla|)f_0^\omega + \frac{\sin(t|\nabla|)}{|\nabla|}f_1^\omega \\ &= \sum_{k \in \mathbb{Z}^d} \epsilon_k \left( \cos(t|\nabla|)f_{0;k}^\omega + \frac{\sin(t|\nabla|)}{|\nabla|}f_{1;k}^\omega \right) \\ (35) \quad &= \sum_{k \in \mathbb{Z}^d} \epsilon_k \left[ \exp(it|\nabla|) \left( \frac{f_{0;k} + i|\nabla|^{-1}f_{1;k}}{2} \right) + \exp(-it|\nabla|) \left( \frac{f_{0;k} - i|\nabla|^{-1}f_{1;k}}{2} \right) \right] \\ &=: \sum_{k \in \mathbb{Z}^d} \epsilon_k \left[ \exp(it|\nabla|)f_k^+ + \exp(-it|\nabla|)f_k^- \right] . \end{aligned}$$

As in (25), we also decompose dyadically in frequency space, and write

$$(36) \quad F_N^\pm := \sum_{\|k\|_\infty \in (N/2, N]} \epsilon_k \exp(\pm it|\nabla|)f_k^\pm \quad \text{and} \quad F_N := F_N^+ + F_N^- .$$

Let  $k \in \mathbb{Z}^d$  with  $\|k\|_\infty \in (N/2, N]$ , and let  $l \in \mathbb{Z}^d$ . We define the tubes  $T_{k,l}^\pm$  by

$$(37) \quad T_{k,l}^+ := \{(t, x) \in [0, N] \times \mathbb{R}^d : \|x - (l \mp t \cdot k / \|k\|_2)\|_2 \leq 1\} .$$

Here, the superscripts  $\pm$  are chosen so that the tubes correspond to the operators  $\exp(\pm it|\nabla|)$ . The dimensions and the shape of the tubes are illustrated in the introduction, see Figure 2. Motivated by the Doppler effect, the tubes  $T_{k,l}^+$  are sometimes called red tubes, and the tubes  $T_{k,l}^-$  are sometimes called blue tubes.

**Proposition 4.1** (Spatial wave packet decomposition).

Let  $k \in \mathbb{Z}^d$  with  $\|k\|_\infty \in (N/2, N]$ . Let  $f_k \in L_x^2(\mathbb{R}^d)$  be a function such that  $\text{supp } \widehat{f}_k \subseteq k + [-1, 1]^d$ . Then,

there exists a decomposition

$$f_k = \sum_{l \in \mathbb{Z}^d} f_{k,l}$$

such that

- (i)  $\text{supp } \widehat{f_{k,l}} \subseteq k + [-4, 4]^d$  for all  $l \in \mathbb{Z}$ ,
- (ii) the family  $\{f_{k,l}\}_{l \in \mathbb{Z}^d}$  satisfies the almost-orthogonality condition

$$(38) \quad \sum_{l \in \mathbb{Z}^d} \|f_{k,l}\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim \|f\|_{L_x^2(\mathbb{R}^d)}^2 ,$$

- (iii) and for any  $D \geq 1$ , any  $l \in \mathbb{Z}^d$ , and all  $(t, x) \in [0, N] \times \mathbb{R}^d$ , it holds that

$$(39) \quad |\exp(\pm it|\nabla|)f_{k,l}(x)| \lesssim_D (1 + \text{dist}((t, x), T_{k,l}^\pm))^{-D} \|f_k\|_{L_x^2(\mathbb{R}^d)} .$$

Wave packet decomposition as in Proposition 4.1 have been used extensively in the literature, see e.g. [6, 18, 27, 30, 57] and the survey [51]. We present the details below, but encourage the expert reader to skip ahead to the end of the proof.

*Proof.* We define the fattened projection

$$\tilde{P}_k := \sum_{\|k' - k\|_\infty \leq 2} P_{k'} .$$

Then, it holds that

$$f = \tilde{P}_k f = \sum_{l \in \mathbb{Z}^d} \tilde{P}_k(\varphi_l f) =: \sum_{l \in \mathbb{Z}^d} f_{k,l} .$$

The frequency support condition (i) directly follows from the definition of  $\tilde{P}_k$ . Furthermore, the almost orthogonality (ii) follows from

$$\sum_{l \in \mathbb{Z}^d} \|f_{k,l}\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim \sum_{l \in \mathbb{Z}^d} \|\varphi_l f_k\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim \|f_k\|_{L_x^2(\mathbb{R}^d)}^2 .$$

Thus, it remains to prove the decay estimate (iii). We only treat the operator  $\exp(it|\nabla|)$ , since the proof for  $\exp(-it|\nabla|)$  is similar. If  $N \lesssim 1$ , the estimate is trivial. Thus, we may assume that  $N \gg 1$ . The argument is based on the method of non-stationary phase. For all  $t \in [0, N]$  and  $x \in \mathbb{R}^d$ , we have that

$$\begin{aligned} & \exp(it|\nabla|)f_{k,l}(x) \\ &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \exp(ix\xi + it|\xi|) \tilde{\psi}(\xi - k) (\widehat{\varphi}_l * \widehat{f})(\xi) d\xi \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \exp(ix\xi + it|\xi|) \tilde{\psi}(\xi - k) \exp(-iy(\xi - \eta)) \varphi(y - l) \widehat{f}(\eta) dy d\eta d\xi \\ &= \frac{1}{(2\pi)^d} \exp(ixk + it|k|) \int_{\mathbb{R}^d} K(\eta; t, x) \exp(-i\eta(k - \eta)) \widehat{f}(\eta) d\eta , \end{aligned}$$

where the kernel  $K(\eta; t, x)$  is given by

$$K(\eta; t, x) := \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \exp(i(x - l)\xi + it(|\xi + k| - |k|)) \exp(-iy(\xi + k - \eta)) \tilde{\psi}(\xi) \varphi(y) d\xi dy .$$

Since  $\text{supp } \varphi \subseteq [-1, 1]^d$ , the function  $a(\xi; y, \eta) := \exp(-iy(\xi + k - \eta))\tilde{\psi}(\xi)\varphi(y)$  has uniformly bounded derivatives in  $\xi$ , i.e, we have for all  $\alpha \in \mathbb{N}_0^d$  that  $|\partial_\xi^\alpha a(\xi; y, \eta)| \lesssim_\alpha 1$ . Using the support conditions in the variables  $y$  and  $\eta$ , it thus suffices to prove for all  $a \in C_c^\infty([-2, 2]^d)$  that

$$(40) \quad \left| \int_{\mathbb{R}^d} \exp(i(x-l)\xi + it(|\xi+k| - |k|))a(\xi)d\xi \right| \lesssim_M (1 + |x-l+t \cdot k/\|k\|_2|)^{-M}.$$

Due to the compact support of  $a(\xi)$ , we restrict to  $|\xi| \leq 2$ . The bound for  $|x-l+t \cdot k/\|k\|_2| \lesssim 1$  is trivial. Thus, we may assume that  $|x-l+t \cdot k/\|k\|_2| \gg 1$ . We define the phase function

$$\Phi_k(\xi) = \Phi_k(\xi; t, x, l) = i(x-l)\xi + it(|\xi+k| - |k|).$$

Then, we have that

$$\begin{aligned} \nabla_\xi \Phi_k(\xi) &= \nabla_\xi \left( (x-l)\xi + t(|\xi+k| - |k|) \right) \\ &= x-l + t \frac{\xi+k}{|\xi+k|} \\ &= x-l + t \frac{k}{|k|} + t \left( \frac{|k| - |\xi-k|}{|\xi+k||k|} k + \frac{\xi}{|\xi+k|} \right). \end{aligned}$$

From the assumption  $|t| \leq N$ , it follows that  $\nabla_\xi \Phi_k = x-l + tk/|k| + \mathcal{O}(1)$ . We also write

$$\nabla_\xi \Phi_k(\xi) = x-l + t \frac{k}{|k|} + t \Psi_k \left( \frac{\xi}{|k|} \right) \frac{k}{|k|},$$

where

$$\Psi_k(\nu) := \frac{1 - |\nu - k/|k||}{|\nu + k/|k||} + \frac{\nu}{|k/|k| + \nu}.$$

From rotation invariance, it follows easily that  $|\nabla^\alpha \Psi_k(\nu)| \lesssim_\alpha 1$  for all  $|\nu| \leq 1/10$ , uniformly in  $k$ . This leads to

$$|\nabla^\alpha \Phi_k(\xi)| \lesssim_\alpha \frac{|t|}{|k|^{|\alpha|-1}} \lesssim 1 \quad \text{for all } |\alpha| \geq 2.$$

We then rewrite the integral in (40) as

$$\begin{aligned} &\int_{\mathbb{R}^d} \exp(i(x-l)\xi + it(|\xi+k| - |k|))a(\xi)d\xi \\ &= \int_{\mathbb{R}^d} \left( \left( \frac{(-i)\nabla_\xi \Phi_k(\xi)\nabla_\xi}{|\nabla \Phi_k(\xi)|^2} \right)^M \exp(i\Phi_k(\xi)) \right) \cdot a(\xi)d\xi \\ &= \int_{\mathbb{R}^d} \exp(i\Phi_k(\xi)) \left( \nabla_\xi \cdot \frac{i\nabla_\xi \Phi_k(\xi)}{|\nabla \Phi_k(\xi)|^2} \right)^M a(\xi)d\xi. \end{aligned}$$

The inequality (40) then follows from the bounds on the phase function above.  $\square$

The wave packet decomposition in Proposition 4.1 is valid on the time interval  $[0, N]$ , and the physical localization deteriorates for larger times. When analyzing the linear evolution on an interval of the form

$[t_0, t_0 + N)$ , with  $t_0 \in N\mathbb{N}_0$ , we therefore use the wave packet decomposition of  $\exp(\pm it_0|\nabla|)f_k$ . To state the result, we set

$$T_{k,l;t_0}^\pm := \{(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^d : \|x - (l \mp (t - t_0) \cdot k/\|k\|_2)\|_2 \leq 1\}.$$

**Corollary 4.2** (Time-translated spatial wave packet decomposition).

Let  $k \in \mathbb{Z}^d$  with  $\|k\|_\infty \in (N/2, N]$ , and let  $t_0 \in N\mathbb{N}_0$ . Let  $f_k \in L_x^2(\mathbb{R}^d)$  be a function satisfying  $\text{supp } \widehat{f}_k \subseteq k + [-1, 1]^d$ . Then, there exists a decomposition

$$\exp(\pm it_0|\nabla|)f_k = \sum_{l \in \mathbb{Z}^d} f_{k,l;t_0}^\pm$$

such that

- (i)  $\text{supp } \widehat{f_{k,l;t_0}^\pm} \subseteq k + [-2, 2]^d$  for all  $l \in \mathbb{Z}^d$ ,
- (ii) the family  $\{f_{k,l;t_0}^\pm\}_{l \in \mathbb{Z}^d}$  satisfies the almost-orthogonality condition

$$(41) \quad \sum_{l \in \mathbb{Z}^d} \|f_{k,l;t_0}^\pm\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim \|f\|_{L_x^2(\mathbb{R}^d)}^2,$$

- (iii) and for any  $D \geq 1$ , any  $l \in \mathbb{Z}^d$ , and all  $(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^d$ , it holds that

$$(42) \quad |\exp(\pm i(t - t_0)|\nabla|)f_{k,l;t_0}^\pm(x)| \lesssim_D (1 + \text{dist}((t, x), T_{k,l;t_0}^\pm))^{-D} \|f_k\|_{L_x^2(\mathbb{R}^d)}.$$

*Proof.* We apply Proposition 4.1 to  $\exp(\pm it_0|\nabla|)f_k$ . □

As discussed in the introduction, we now group the wave packets into bushes and a (nearly) non-overlapping collection (see Figure 3). This argument is inspired by Bourgain's bush argument from [6], and we also refer the reader to [56, Proposition 2.2].

Before we state main proposition, we define the truncated and fattened  $\ell^\infty$ -cone

$$(43) \quad \widetilde{K}_{t_0, x_0}^N := \{(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^d : \|x - x_0\|_\infty \leq 16N - |t - t_0|\}.$$

The significance of  $\widetilde{K}_{t_0, x_0}^N$  will be explained in Section 5.2 and Section 6. For now, we encourage the reader to treat  $\widetilde{K}_{t_0, x_0}^N$  as space-time cube of scale  $N$ .

**Proposition 4.3** (Wave packet decomposition and bushes).

Let  $\{f_k^\pm\}_k \subseteq L_x^2(\mathbb{R}^d)$  be a family of functions, where  $\|k\|_\infty \in (N/2, N]$ , and  $\text{supp } \widehat{f}_k^\pm \subseteq k + [-1, 1]^d$ . Let  $t_0 \in N\mathbb{N}_0$ , let  $x_0 \in N\mathbb{Z}^d$ , and let the wave packets  $\{f_{k,l;t_0}^\pm\}$  be as in Corollary 4.2. Furthermore, let  $\mathcal{Q}_{t_0, x_0}^N$  be a collection of disjoint space-time cubes with sidelength  $\sim N^\delta$  covering  $\widetilde{K}_{t_0, x_0}^N$ . We group the wave packets according to their amplitude by setting

$$(44) \quad \mathcal{A}_m = \mathcal{A}_{m, t_0, x_0}^{N, \pm} := \{(k, l) \in \mathbb{Z}^d \times \mathbb{Z}^d : \|f_{k,l;t_0}^\pm\|_{L_x^2(\mathbb{R}^d)} \in [2^m, 2^{m+1}], \|l - x_0\|_\infty \leq 3N\}.$$

Then, there exists a family of bushes  $\{\mathcal{B}_{j,m}\}_j = \{\mathcal{B}_{j,m,t_0,x_0}^{N,\pm}\}_j$ , where  $j = 1, \dots, J_{m,t_0,x_0}^{N,\pm}$ , and a nearly non-overlapping set  $\mathcal{D}_m = \mathcal{D}_{m,t_0,x_0}^{N,\pm}$ , depending only on the set  $\mathcal{A}_m$ , so that the following holds:

(i) The sets form a partition of  $\mathcal{A}_m$ , i.e.,

$$(45) \quad \mathcal{A}_m = \mathcal{D}_m \dot{\cup} \left( \dot{\cup}_{j=1, \dots, J} \mathcal{B}_{j,m} \right)$$

(ii) We have the bound on the number of wave packets

$$(46) \quad \sum_{x_0 \in N\mathbb{Z}^d} \sum_{m \in \mathbb{Z}} 2^{2m} \# \mathcal{A}_{m,t_0,x_0}^{N,\pm} \lesssim \sum_k \|f_k\|_{L_x^2(\mathbb{R}^d)}^2 .$$

(iii) Each bush  $\mathcal{B}_{j,m}$  contains at least  $\mu = \mu(N, m) := N^{-\frac{1}{2}} \# \mathcal{A}_m$  wave packets.

(iv) For each bush  $\mathcal{B}_{j,m}$ , all corresponding wave packets intersect in the same region of space-time. More precisely, there exists a cube  $Q \in \mathcal{Q}_{t_0,x_0}^N$  s.t.

$$(47) \quad T_{k,l;t_0}^\pm \cap 2Q \neq \emptyset \quad \forall (k,l) \in \mathcal{B}_{j,m} .$$

(v) At  $\mu = N^{-\frac{1}{2}} \# \mathcal{A}_m$  wave packets in  $\mathcal{D}_m$  overlap, i.e., we have for all cubes  $Q \in \mathcal{Q}_{t_0,x_0}^N$  that

$$(48) \quad \#\{(k,l) \in \mathcal{D}_m : T_{k,l;t_0}^\pm \cap 2Q \neq \emptyset\} < P .$$

The choice of the number of packets/multiplicity  $\mu = N^{-\frac{1}{2}} \# \mathcal{A}_m$  will be justified in the proof of Proposition 6.1, see (65). The parameter  $\mu$  corresponds to the multiplicity parameter in Bourgain's bush argument, see [56, Proposition 2.2].

**Remark 4.4.**

We will later apply this proposition to a set of random functions  $\{\epsilon_k f_k^\pm\}_k$ . From (44), it follows that the sets  $\mathcal{A}_m$ , and hence also  $\mathcal{B}_{j,m}$  and  $\mathcal{D}_m$ , do not depend on the random signs  $\{\epsilon_k\}$ .

*Proof.* Let us first prove the inequality in (ii). From Corollary 4.2, it follows that

$$\sum_{x_0 \in N\mathbb{Z}^d} \sum_{m \in \mathbb{Z}} 2^{2m} \# \mathcal{A}_{m,t_0,x_0}^{N,\pm} \lesssim \sum_{\substack{k,l \in \mathbb{Z}^d \\ \|k\|_\infty \in (N/2, N]}} \|f_{k,l;t_0}^\pm\|_{L_x^2}^2 \lesssim \sum_{\substack{k \in \mathbb{Z}^d \\ \|k\|_\infty \in (N/2, N]}} \|f_k^\pm\|_{L_x^2}^2 .$$

We now construct the sets  $\mathcal{B}_{j,m,t_0,x_0}^{N,\pm}$  and  $\mathcal{D}_{m,t_0,x_0}^{N,\pm}$ . To simplify the expressions, we drop the super- and subscripts  $\pm, N, t_0$ , and  $x_0$  from our notation. The basic idea is to form the bushes through a greedy selection algorithm. For any  $Q \in \mathcal{Q}_{t_0,x_0}^N$ , we define

$$(49) \quad \mathcal{T}_m(Q) := \{(k,l) \in \mathcal{A}_m : T_{k,l} \cap Q \neq \emptyset\} .$$

We further set  $\mathcal{T}_m^{(1)}(Q) := \mathcal{T}_m(Q)$ . We then choose a cube  $Q_1 \in \mathcal{Q}_{t_0,x_0}^N$  such that

$$\#\mathcal{T}_m^{(1)}(Q_1) = \max_{Q \in \mathcal{Q}_{t_0,x_0}^N} \#\mathcal{T}_m^{(1)}(Q) ,$$

and define the first bush as  $\mathcal{B}_{1,m} := \mathcal{T}_m^{(1)}(Q_1)$ . By setting

$$\mathcal{T}_m^{(2)}(Q) = \mathcal{T}_m^{(1)}(Q) \setminus \mathcal{B}_{1,m} ,$$

we remove all of the wave packets in the first bush from the collection. We then iteratively define  $\mathcal{B}_{j,m} := \mathcal{T}_m^{(j)}(Q_j)$ , where

$$\#\mathcal{T}_m^{(j)}(Q_j) = \max_{Q \in \mathcal{Q}_{t_0, x_0}^N} \#\mathcal{T}_m^{(j)}(Q),$$

and the collections  $\mathcal{T}_m^{(j)}(Q)$  are defined as  $\mathcal{T}_m^{(j-1)}(Q) \setminus \mathcal{B}_{j-1,m}$ . Once  $\mathcal{T}_m^{(j+1)}(Q_{j+1}) < \mu$ , we no longer create a new bush, and instead stop the algorithm. Since  $\mathcal{A}_m$  contains at most  $\sim N^8$  wave packets, the greedy selection algorithm terminates after finitely many steps. From the construction, we see that the sets  $\mathcal{B}_{j,m} \subseteq \mathbb{Z}^d \times \mathbb{Z}^d$  are disjoint (even though the corresponding tubes may still overlap). Finally, we define the collection  $\mathcal{D}_m$  by

$$\mathcal{D}_m := \mathcal{A}_m \setminus \bigcup_{j=1}^J \mathcal{B}_{j,m}.$$

The properties (i), (iii), (iv), and (v) then follow directly from the construction.  $\square$

We now prove a probabilistic estimate for the wave packets with random coefficients.

**Proposition 4.5** (Square-root cancellation for wave packets).

Let  $\{f_k^\pm\}_k \subseteq L_x^2(\mathbb{R}^d)$  be a deterministic family of functions, where  $k \in \mathbb{Z}^d$  satisfies  $\|k\|_\infty \in (N/2, N]$ , and assume that  $\sum_k \|f_k^\pm\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim 1$ . Let  $\theta > 0$  be a parameter, and let  $C_d > 0$  be any large absolute constant. Then, we have for all  $m \in \mathbb{Z}$  satisfying  $-C_d \log(N) \leq m \leq C_d \log(N)$  that

$$(50) \quad \mathbb{E} \left[ \sup_{\substack{t_0=0, \dots, \lfloor N^\theta \rfloor N \\ x_0 \in N\mathbb{Z}^d}} \sup_{j=1, \dots, J_{m, t_n, x_0}^{N, \pm}} \frac{\left\| \sum_{(k,l) \in \mathcal{B}_{j,m,t_0,x_0}^{N,\pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_{t,x}^\infty(\mathbb{R} \times \mathbb{R}^d)} \right] \lesssim N^{\delta d}$$

and

$$(51) \quad \mathbb{E} \left[ \sup_{\substack{t_0=0, \dots, \lfloor N^\theta \rfloor N \\ x_0 \in N\mathbb{Z}^d}} \frac{\left\| \sum_{(k,l) \in \mathcal{D}_{m,t_0,x_0}^{N,\pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^\infty L_x^\infty([t_0, t_0+N] \times \mathbb{R}^d)} \right] \lesssim N^{\delta d}.$$

Here,  $\mu = N^{-\frac{1}{2}} \#\mathcal{A}_{m,t_0,x_0}^{N,\pm}$  is as in Proposition 4.3. To be perfectly precise, we use the convention  $0/0 := 0$  in (51).

The expressions in (50) and (51) may seem complicated. To make sense of them, recall that the square function heuristic predicts that  $\sum_k \epsilon_k a_k$  is roughly of size  $\sim (\sum_k a_k^2)^{\frac{1}{2}}$ . Then, Proposition 4.5 simply states that the square function heuristic can be justified for all relevant amplitudes, for all relevant times, all positions, all families of bushes, and all non-overlapping collections.

For instance, let us heuristically motivate (51). By the definition of  $\mathcal{D}_{m,t_0,x_0}^{N,\pm}$ , any fixed point in the space-time region  $[t_0, t_0+N] \times \mathbb{R}^d$  is contained in the (moral) support of at most  $\mu$  wave packets. Since each of the wave packets has amplitude  $\sim 2^m$ , and they all correspond to different frequencies  $k \in \mathbb{Z}^d$ , the square-function heuristic predicts a contribution of size  $\sim 2^m \mu^{\frac{1}{2}}$ .

*Proof.* In this proof, we make extensive use of Lemma 2.4. First, we prove that the suprema in (50) and (51) are over at most  $N^{\mathcal{O}(C_d)}$ -many terms. From (46), it follows for all  $m \geq -C_d \log(N)$  that

$$\sum_{x_0 \in N\mathbb{Z}^d} \#\mathcal{A}_{m,t_0,x_0}^{N,\pm} \lesssim 2^{-2m} \sum_{k,l} \|f_{k,l}^\pm\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim 2^{-2m} \sum_k \|f_k^\pm\|_{L_x^2(\mathbb{R}^d)}^2 \lesssim N^{2C_d}.$$

Thus, this bounds the number of all wave packets with amplitude  $\sim 2^m$ . Since each bush  $\mathcal{B}_{j,m,t_0,x_0}^{N,\pm}$  contains at least one wave packet, the supremum in (50) is over at most  $\mathcal{O}(N^{2C_d})$  non-zero terms. The same applies to the non-overlapping families  $\mathcal{D}_{m,t_0,x_0}^{N,\pm}$  in (51). From Lemma 2.4, it then suffices to obtain uniform sub-gaussian bounds on each individual term in (50) and (51).

We start with the contribution of the bushes. To simplify the notation, we write  $\mathcal{B}_{j,m} = \mathcal{B}_{j,m,t_0,x_0}^{N,\pm}$ . From Bernstein's inequality and Lemma 2.6, we have for all  $2 \leq p < \infty$  that

$$\begin{aligned} & \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R}^d)} \\ & \leq N^{\frac{d+1}{p}} \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{L_t^p L_x^p(\mathbb{R} \times \mathbb{R}^d)}. \end{aligned}$$

Before we utilize the randomness, we observe that for each  $k \in \mathbb{Z}^d$  at most  $\mathcal{O}(N^{\delta d})$  tubes  $T_{k,l;t_0}^\pm$  can intersect a space-time cube of sidelength  $\sim N^\delta$ . As a result, it follows from (47) that

$$\#\{l \in \mathbb{Z}^d : (k,l) \in \mathcal{B}_{j,m}\} \lesssim N^{\delta d}.$$

For all  $p \leq r < \infty$ , we then obtain from Minkowski's integral inequality, Khintchine's inequality, and the refined Strichartz estimate (Corollary 2.13) that

$$\begin{aligned} & \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{L_\omega^\infty L_t^\infty L_x^\infty(\Omega \times \mathbb{R} \times \mathbb{R}^d)} \\ & \lesssim N^{\frac{d+1}{p}} \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{L_t^p L_x^p L_\omega^r(\mathbb{R} \times \mathbb{R}^d \times \Omega)} \\ & \lesssim \sqrt{r} N^{\frac{d+1}{p}} \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{B}_{j,m}} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^p L_x^p(\mathbb{R} \times \mathbb{R}^d)} \\ & \lesssim \sqrt{r} N^{\frac{d+1}{p} + \frac{\delta d}{2}} \left\| \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{L_t^p L_x^p \ell_{k,l}^2(\mathbb{R} \times \mathbb{R}^d \times \mathcal{B}_{j,m})} \\ & \lesssim \sqrt{r} N^{\frac{d+1}{p} + \frac{\delta d}{2}} \left\| \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^+ \right\|_{\ell_{k,l}^2 L_t^p L_x^p(\mathcal{B}_{j,m} \times \mathbb{R} \times \mathbb{R}^d)} \\ & \lesssim \sqrt{r} N^{\frac{d+2}{p} + \frac{\delta d}{2}} \left\| f_{k,l;t_0}^+ \right\|_{\ell_{k,l}^2 L_x^2(\mathcal{B}_{j,m} \times \mathbb{R}^d)} \\ & \lesssim \sqrt{r} N^{\frac{d+2}{p} + \frac{\delta d}{2}} 2^m (\#\mathcal{B}_{j,m})^{\frac{1}{2}}. \end{aligned}$$

By taking  $p \geq 2$  to be sufficiently large, we then obtain the desired sub-gaussian bound. This completes the proof of (50).



We now control the contribution of a single non-overlapping family  $\mathcal{D}_m = \mathcal{D}_{m,t_0,x_0}^{N,\pm}$ . For the technical aspects of this part, recall that the collection  $\mathcal{Q}_{t_0,x_0}^N$  from Proposition 4.3 covers  $\tilde{K}_{t_0,x_0}^N$ , but due the definition of  $\mathcal{A}_m$  in (44), all the tubes  $T_{k,l;t_0}^\pm$  with indices in  $\mathcal{D}_m$  are contained in the region  $\|x-x_0\|_\infty \leq 6N$ . This gives us sufficient room for the following argument.

We let  $2 \leq p < \infty$ . As before, it follows from Bernstein's inequality and Lemma 2.6 that

$$\begin{aligned} & \left\| \sum_{(k,l) \in \mathcal{D}_{m,t_0,x_0}^{N,\pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^\infty L_x^\infty([t_0, t_0+N] \times \mathbb{R}^d)} \\ & \lesssim N^{\frac{d+1}{p}} \left\| \sum_{(k,l) \in \mathcal{D}_{m,t_0,x_0}^{N,\pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^p L_x^p([t_0, t_0+N] \times \mathbb{R}^d)} \end{aligned}$$

For all  $p \leq r < \infty$ , we then obtain from Minkowski's integral inequality and Khintchine's inequality that

$$\begin{aligned} & \left\| \sum_{(k,l) \in \mathcal{D}_m} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_\omega^r L_t^\infty L_x^\infty(\Omega \times [t_0, t_0+N] \times \mathbb{R}^d)} \\ & \lesssim N^{\frac{d+1}{p}} \left\| \sum_{(k,l) \in \mathcal{D}_m} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^p L_x^p L_\omega^r([t_0, t_0+N] \times \mathbb{R}^d \times \Omega)} \\ & \lesssim \sqrt{r} N^{\frac{d+1}{p}} \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^p L_x^p([t_0, t_0+N] \times \mathbb{R}^d)} \\ & \lesssim \sqrt{r} N^{\frac{d+1}{p}} \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^p L_x^p(\tilde{K}_{t_0,x_0}^N)} \\ & \quad + \sqrt{r} N^{\frac{d+1}{p}} \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^p L_x^p([t_0, t_0+N] \times \mathbb{R}^d) \setminus \tilde{K}_{t_0,x_0}^N} \end{aligned}$$

Since  $\mu = N^{-\frac{1}{2}} \#\mathcal{A}_m \geq N^{-\frac{1}{2}}$ , the bound on  $([t_0, t_0+N] \times \mathbb{R}^d) \setminus \tilde{K}_{t_0,x_0}^N$  easily follows from the decay estimate (42). Thus, we now control the contribution on  $\tilde{K}_{t_0,x_0}^N$ . From Hölder's inequality, we have that

$$\begin{aligned} & \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^p L_x^p(\tilde{K}_{t_0,x_0}^N)} \\ & \lesssim N^{\frac{d+1}{p}} \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^\infty L_x^\infty(\tilde{K}_{t_0,x_0}^N)} \end{aligned}$$

Now pick any cube  $Q \in \mathcal{Q}_{t_0,x_0}^N$ . In analogy to (49), we define the collection of "remaining" tubes by

$$\mathcal{T}_m^r(Q) := \{(k,l) \in \mathcal{D}_m : T_{k,l;t_0}^\pm \cap 2Q \neq \emptyset\}.$$

From Proposition 4.3, it follows that  $\#\mathcal{T}_m^r(Q) \leq \mu$ . As above, we have for each frequency  $k \in \mathbb{Z}^d$  the bound  $\#\{l \in \mathbb{Z}^d : (k,l) \in \mathcal{T}_m^r(Q)\} \lesssim N^{\delta d}$ . Using the decay estimate (49) to treat distant wave packets, we

obtain

$$\begin{aligned}
& \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^\infty L_x^\infty(Q)} \\
& \lesssim \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{T}_m^r(Q)} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^\infty L_x^\infty(Q)} \\
& \quad + \left\| \left( \sum_{k \in \mathbb{Z}^d} \left( \sum_{l: (k,l) \in \mathcal{D}_m \setminus \mathcal{T}_m^r(Q)} \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right)^2 \right)^{\frac{1}{2}} \right\|_{L_t^\infty L_x^\infty(Q)} \\
& \lesssim N^{\frac{\delta d}{2}} \left\| \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^\infty L_x^\infty \ell_{k,l}^2(Q \times \mathcal{T}_m^r(Q))} + N^{-10C_d} \\
& \lesssim N^{\frac{\delta d}{2}} 2^m (\#\mathcal{T}_m^r(Q))^{\frac{1}{2}} + N^{-10C_d} \\
& \lesssim N^{\frac{\delta d}{2}} 2^m \mu^{\frac{1}{2}}.
\end{aligned}$$

After taking the supremum over all cubes  $Q \in \mathcal{Q}_{t_0, x_0}^N$ , we finally arrive at

$$\left\| \sum_{(k,l) \in \mathcal{D}_m} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_\omega L_t^\infty L_x^\infty(\Omega \times [t_0, t_0+N] \times \mathbb{R}^d)} \lesssim \sqrt{r} N^{2\frac{d+2}{p} + \frac{\delta d}{2}} 2^m \mu^{\frac{1}{2}}.$$

By choosing  $p \geq 2$  sufficiently large, we arrive at the desired sub-gaussian bound.  $\square$

**Definition 4.6** (Wave packet “norm”).

Let  $(f_0, f_1) \in H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$  and let  $f_k^\pm$  be defined as in (34) and (35). For any  $N_0 \geq 1$ , we then define the wave packet “norm” of the random data  $(f_0^\omega, f_1^\omega)$  by

$$\begin{aligned}
& \|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_0)} := \\
& \sum_{\substack{N \geq N_0 \\ |m| \leq C_d \log(N)}} N^{-2\delta d} \sup_{\substack{t_0=0, \dots, [N^\theta]N \\ x_0 \in N\mathbb{Z}^d}} \sup_{\substack{j=1, \dots, J_{m, t_0, x_0}^{N, \pm} \\ x_0 \in N\mathbb{Z}^d}} \frac{\left\| \sum_{(k,l) \in \mathcal{B}_{j, m, t_0, x_0}^{N, \pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R}^d)}}{2^m (\#\mathcal{B}_{j, m, t_0, x_0}^{N, \pm})^{\frac{1}{2}}} \\
& + \sum_{\substack{N \geq N_0 \\ |m| \leq C_d \log(N)}} N^{-2\delta d} \sup_{\substack{t_0=0, \dots, [N^\theta]N \\ x_0 \in N\mathbb{Z}^d}} \frac{\left\| \sum_{(k,l) \in \mathcal{D}_{m, t_0, x_0}^{N, \pm}} \epsilon_k \exp(\pm i(t-t_0)|\nabla|) f_{k,l;t_0}^\pm \right\|_{L_t^\infty L_x^\infty([t_0, t_0+N] \times \mathbb{R}^d)}}{2^m \mu^{\frac{1}{2}}}
\end{aligned}$$

While the quantity  $\|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_0)}$  measures the size of the wave packets (over their expected size), it is certainly far from being an actual norm.

**Corollary 4.7.**

Let  $(f_0, f_1) \in H^s(\mathbb{R}^d) \times H^{s-1}(\mathbb{R}^d)$  and let  $f_k^\pm$  be defined as in (34) and (35). Furthermore, we assume that  $N_0 = N_0(\{Y_{k,l}\})$  satisfies

$$\sum_{\|k\|_\infty \geq N_0/2} \left( \|f_0; k\|_{L_x^2(\mathbb{R}^d)}^2 + \|f_1; k\|_{\dot{H}_x^{-1}(\mathbb{R}^d)}^2 \right) \lesssim 1.$$

Then, it holds that

$$\mathbb{E}_\epsilon \|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_0)} < \infty ,$$

where  $\mathbb{E}_\epsilon$  denotes the expectation over the random signs  $\{\epsilon_k\}$ .

*Proof.* This follows directly from Proposition 4.5 and Definition 4.6.  $\square$

For the bootstrap argument in Section 6, it will be convenient to create a small forcing term by truncating to high frequencies. If  $N_{\text{hi}} = N_{\text{hi}}(\omega)$  is a (possibly random) frequency parameter, we set

$$(52) \quad F_{\text{hi}} := \sum_{N \geq N_{\text{hi}}} \left( \cos(t|\nabla|) f_{0,N}^\omega + \frac{\sin(t|\nabla|)}{|\nabla|} f_{1,N}^\omega \right) .$$

**Proposition 4.8** (Truncation to high frequencies).

Let  $0 < \eta \leq 1$  be an absolute constant and let  $s > \frac{1}{3}$ . Let  $(f_0, f_1) \in H_x^s(\mathbb{R}^4) \times H_x^{s-1}(\mathbb{R}^4)$ , let  $(f_0^\omega, f_1^\omega)$  be their microlocal randomizations, and let  $\{f_k^\pm\}$  be as in (34) and (35). Then, there exists a random frequency parameter  $N_{\text{hi}} \geq \eta^{-1}$  such that

$$\begin{aligned} \|(P_{\geq N_{\text{hi}}/4} f_0^\omega, P_{\geq N_{\text{hi}}/4} f_1^\omega)\|_{H^s \times H^{s-1}} &\leq \eta , \\ \|(f_0^\omega, f_1^\omega)\|_{Z(N_{\text{hi}})} &\leq \eta , \\ \|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_{\text{hi}})} &\leq \eta , \\ \|F_{\text{hi}}\|_{L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)} &\leq \eta . \end{aligned}$$

*Proof.* We only need to combine the previous estimates. From Lemma 2.11, it follows that

$$\sum_{N \geq 2} \sum_{\|k\|_\infty \in (N/2, N]} \left( N^{2s} \|f_{0;k}\|_{L_x^2}^2 + N^{2(s-1)} \|f_{1;k}\|_{L_x^2}^2 \right) < \infty \quad \text{a.s.}$$

From dominated convergence, it then follows that there exists some random frequency  $N_{\text{hi},1}$ , depending only on the random variables  $\{Y_{k,l}\}$ , satisfying

$$\sum_{N \geq N_{\text{hi},1}} \sum_{\|k\|_\infty \in (N/2, N]} \left( N^{2s} \|f_{0;k}\|_{L_x^2}^2 + N^{2(s-1)} \|f_{1;k}\|_{L_x^2}^2 \right) \leq \eta \quad \text{a.s.}$$

From Corollary 4.7, it then follows that

$$\mathbb{E}_\epsilon \|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_{\text{hi},1})} < \infty .$$

From dominated convergence, it follows that there exists a random frequency  $N_{\text{hi},2}$ , depending on the random variables  $\{\epsilon_k\}$  and  $\{Y_{k,l}\}$ , which satisfies

$$\|(f_0^\omega, f_1^\omega)\|_{\text{WP}(N_{\text{hi},2})} \leq \eta \quad \text{a.s.}$$

By similar arguments, it also follows from Proposition 3.1 and Corollary 3.3 that there exists random frequencies  $N_{\text{hi},3}$  and  $N_{\text{hi},4}$  such that

$$\|(f_0^\omega, f_1^\omega)\|_{Z(N_{\text{hi},3})} \leq \eta \quad \text{and} \quad \sum_{N \geq N_{\text{hi},4}} \left\| \cos(t|\nabla|) f_{0,N}^\omega + \frac{\sin(t|\nabla|)}{|\nabla|} f_{1,N}^\omega \right\|_{L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)} \leq \eta$$

For the second inequality, we have used the condition  $s > \frac{1}{3}$ .

By choosing  $N_{\text{hi}} := \max(4N_{\text{hi},1}, N_{\text{hi},2}, N_{\text{hi},3}, N_{\text{hi},4}, \eta^{-1})$ , we arrive at the desired conclusion.  $\square$

## 5. NONLINEAR EVOLUTION: LOCAL WELL-POSEDNESS, STABILITY THEORY, AND FLUX ESTIMATES

In this section, we first apply to Da Prato-Debussche trick [19] to the nonlinear wave equation with random initial data. Then, we recall certain properties of the (forced) energy critical nonlinear wave equation. In our exposition of the local well-posedness and stability theory, we mainly rely on [26]. The flux estimate already played a major role in the author's work on almost sure scattering for the radial energy critical NLW [12], and we loosely follow parts of [12, Section 6].

Let  $N_{\text{hi}}$  be as in Proposition 4.8, and let  $F = F_{\text{hi}}$  be as in (52). We then decompose the solution  $u$  of (5) by setting  $v := u - F$ . Then, the nonlinear component  $v$  solves the forced nonlinear wave equation

$$(53) \quad \begin{cases} -\partial_{tt}v + \Delta v = (v + F)^3 & (t, x) \in \mathbb{R} \times \mathbb{R}^4, \\ v|_{t=0} = v_0 \in \dot{H}^1, \quad \partial_t v|_{t=0} = v_1 \in L^2, \end{cases}$$

where  $v_0 := u_0 + f_{0, < N_{\text{hi}}}^\omega$  and  $v_1 := u_1 + f_{1, < N_{\text{hi}}}^\omega$ . The randomness in the initial data  $(v_0, v_1)$  is not important, and we treat it as arbitrary data in the energy space. For the rest of this section, we treat  $F$  as an arbitrary forcing term in  $L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)$ , since the finer properties of  $F$  will only be relevant in Section 6.

**5.1. Local well-posedness and stability theory.** In this section, we recall the local well-posedness of (53). Using stability theory, we recall the reduction of Theorem 1.3 to an a-priori energy bound. These results are well-known in the literature, see e.g. [26, 44].

**Lemma 5.1** (Local well-posedness [26, Lemma 3.1]).

Let  $(v_0, v_1) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  and  $F \in L_t^3 L_x^6([0, \infty) \times \mathbb{R}^4)$ . Then, there exists a time  $T > 0$  and a unique solution  $v: [0, T) \times \mathbb{R}^4 \rightarrow \mathbb{R}$  satisfying

$$v \in C_t^0 \dot{H}_x^1([0, T) \times \mathbb{R}^4) \cap L_t^3 L_x^6([0, T) \times \mathbb{R}^4) \quad \text{and} \quad \partial_t v \in C_t^0 L_x^2([0, T) \times \mathbb{R}^4).$$

Using stability theory, [26] proved the following proposition.

**Proposition 5.2** (Reduction to an a-priori energy bound [26, Theorem 1.3]).

Let  $(v_0, v_1) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  and  $F \in L_t^3 L_x^6(\mathbb{R} \times \mathbb{R}^4)$ . Let  $v$  be a solution of (53), and let  $T_+ > 0$  be its maximal time of existence. Furthermore, assume the a-priori energy bound

$$\sup_{t \in [0, T_+)} E[v](t) < \infty.$$

Then,  $v$  is a global solution and satisfies the global space-time bound  $\|v\|_{L_t^3 L_x^6([0, \infty) \times \mathbb{R}^4)} < \infty$ . As a result, there exist scattering states  $(v_0^+, v_1^+) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$  such that

$$\lim_{t \rightarrow +\infty} \|(v(t) - W(t)(v_0^+, v_1^+), \partial_t v(t) - \partial_t W(t)(v_0^+, v_1^+))\|_{\dot{H}^1 \times L^2} = 0.$$

Using Lemma 5.1 and Proposition 5.2, we have reduced the proof of Theorem 1.3 to an a-priori bound on the energy of  $v$ .

**5.2. Flux estimates.** As before, we let  $v: I \times \mathbb{R}^4 \rightarrow \mathbb{R}$  be a solution to the forced equation (53). Recall that the (symmetric) energy-momentum tensor for the energy critical nonlinear wave equation is given by

$$\begin{aligned} T^{00} &= T^{00}[v] := \frac{1}{2}((\partial_t v)^2 + |\nabla v|^2) + \frac{1}{4}v^4, \\ T^{j0} &= T^{j0}[v] := -\partial_t v \cdot \partial_{x_j} v, \\ T^{jk} &= T^{jk}[v] := \partial_{x_j} v \partial_{x_k} v - \frac{\delta^{jk}}{4}(-\partial_{tt} + \Delta)(v^2) + \frac{\delta^{jk}}{2}v^4. \end{aligned}$$

The component  $T^{00}$  is the energy density, the component  $T^{j0}$  is the  $j$ -th momentum/energy flux, and the components  $T^{jk}$  are called the momentum flux. If  $v$  solves the energy critical nonlinear wave equation (1), then the energy-momentum tensor is divergence free. This fails for solutions to the forced equation (53); however, one can still expect that the error terms have lower order. Setting  $\mathcal{N} = (v + F)^3 - v^3$ , a short computation shows that

$$(54) \quad \partial_t T^{00} + \partial_{x_k} T^{0k} = -\mathcal{N} \partial_t v,$$

$$(55) \quad \partial_t T^{j0} + \partial_{x_k} T^{jk} = \mathcal{N} \partial_{x_j} v - \frac{1}{2} \partial_{x_j} (\mathcal{N} v).$$

As in earlier work on almost sure scattering for radial data [12, 26, 25, 32], the main goal of this paper is to bound the energy of  $v$ . In terms of the energy-momentum tensor, the (total) energy can be written as

$$E[v](t) := \int_{\mathbb{R}^4} T^{00}(t, x) dx.$$

For future use, we record the following consequence of (54).

**Lemma 5.3** (Total energy increment).

Let  $v: I \times \mathbb{R}^4 \rightarrow \mathbb{R}$  be a solution of (53), and let  $a, b \in I$  with  $a \leq b$ . Then, we have that

$$(56) \quad E(b) - E(a) = - \int_a^b \int_{\mathbb{R}^4} \mathcal{N} \partial_t v dx dt \leq 6 \int_a^b \int_{\mathbb{R}^4} |F| |v|^2 |\partial_t v| dx dt + 3 \int_a^b \int_{\mathbb{R}^4} |F|^3 |\partial_t v| dx dt.$$

We will later see that the second summand on the right-hand side of (56) can be bounded directly using Hölder's inequality and probabilistic Strichartz estimates. In contrast, no such estimate is available for the first summand, and we need the wave packet decomposition to control this term.

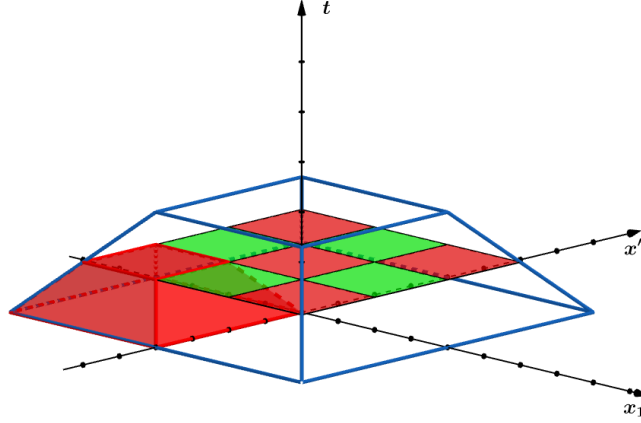
Once we employ the wave packet decomposition, it will be natural to study the energy on a time and length scale  $\sim N \geq 1$ . We fix  $t_0 \in N\mathbb{N}_0$  and  $x_0 \in NZ^4$ , and define the local energy

$$(57) \quad E_{t_0, x_0}^N[v](t) := \int_{\|x - x_0\|_\infty \leq 2N - |t - t_0|} T^{00}(t, x) dx, \quad \text{where } t \in [t_0, t_0 + N].$$

Thus, this definition is adapted to the truncated  $\ell^\infty$ -cone  $K_{t_0, x_0}^N$ , which is given by

$$(58) \quad K_{t_0, x_0}^N := \{(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^4 : \|x - x_0\|_\infty \leq 2N - |t - t_0|\}.$$

It might be more appropriate to call  $K_{t_0, x_0}^N$  a pyramid (see Figure 5); unfortunately, the letter  $P$  is already heavily used in our notation, so that we decided to use the letter  $K$ . Our reason for using the  $\ell^\infty$ -norm, instead of the more common  $\ell^2$ -norm, lies in the induction on scales argument (Proposition 56). Then, it will be an advantage to write  $K_{t_0, x_0}^N$  as the union of finitely overlapping smaller cones  $K_{\tau_0, y_0}^M$ , which are



In this figure, we illustrate the truncated  $\ell^\infty$ -cone, and its decomposition into smaller cones. The blue lines are the edges of a large  $\ell^\infty$ -cone. The red and green squares are the tops of smaller truncated cones. In the lower left corner, we have drawn a single one of these smaller cones. As can be seen from this figure, no smaller cone has to exit the large cone.

FIGURE 5. Decomposition of the  $\ell^\infty$ -cone

contained in  $K_{t_0, x_0}^N$ . Using finite speed of propagation and the inequality  $\|\cdot\|_{\ell^\infty(\mathbb{R}^4)} \leq \|\cdot\|_{\ell^2(\mathbb{R}^4)}$ , one can still meaningfully restrict the nonlinear wave equation to  $K_{t_0, x_0}^N$ .

**Lemma 5.4** (Local energy increment).

Let  $v$  be a solution to the forced equation (53), let  $N \geq 1$ , let  $t_0 \in N\mathbb{N}_0$ , and let  $x_0 \in NZ^4$ . Then, we have that

$$(59) \quad \sup_{t \in [t_0, t_0 + N]} E_{t_0, x_0}^N[v](t) \leq E_{t_0, x_0}^N[v](t_0) + 6 \int_{K_{t_0, x_0}^N} |F||v|^2 |\partial_t v| dx dt + 3 \int_{K_{t_0, x_0}^N} |F|^3 |\partial_t v| dx dt .$$

*Proof.* Using (54), we have that

$$\begin{aligned} \frac{d}{dt} E_{t_0, x_0}^N[v](t) &= - \int_{\substack{\|x-x_0\|_\infty \\ = 2N-|t-t_0|}} T^{00} d\sigma(x) + \int_{\substack{\|x-x_0\|_\infty \\ \leq 2N-|t-t_0|}} \partial_t T^{00} dx \\ &= \int_{\substack{\|x-x_0\|_\infty \\ = 2N-|t-t_0|}} (-T^{00} + T^{0j} \nu_j) d\sigma(x) - \int_{\substack{\|x-x_0\|_\infty \\ \leq 2N-|t-t_0|}} \mathcal{N} \partial_t v dx \end{aligned}$$

Here,  $\nu$  is the outward unit normal to the cube. From Cauchy-Schwarz, it follows that  $|T^{0j} \nu_j| \leq |\partial_t v| \|\nabla v\|_2 \leq T^{00}$ . After integrating in time, this completes the proof.  $\square$

To simplify the notation, we now write

$$(60) \quad \mathcal{E}_{t_0, x_0}^N[v] := \sup_{t \in [t_0, t_0 + N]} E_{t_0, x_0}^N[v](t) .$$

In the following, we want to deduce a flux estimate for the solution  $v$  of the forced NLW (53). Here, we encounter a minor technical problem. Let  $(t', x') \in K_{t_0, x_0}^N$  be a point in the truncated  $\ell^\infty$ -cone. We then want to control the potential  $|v|^4$  on the truncated light-cone

$$C_{t', x'}^N := \{(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^4 : |t - t'| = \|x - x'\|_2\} .$$

Unfortunately,  $C_{t', x'}^N$  may not be contained in  $K_{t_0, x_0}^N$ , and hence we cannot expect to bound this solely by  $\mathcal{E}_{t_0, x_0}^N[v]$ . Since the flux estimate is derived through a monotonicity formula for the local energy, this issue persists even if we are only interested in the portion of  $C_{t', x'}^N$  intersecting  $K_{t_0, x_0}^N$ . To solve this problem, while still keeping the same energy increment as in (59), we introduce the notion of a locally forced solution.

**Definition 5.5** (Locally forced solution).

Let  $t_0 \in N\mathbb{N}_0$  and  $x_0 \in N\mathbb{Z}^4$ . We call  $w : \mathbb{R} \times \mathbb{R}^4 \rightarrow \mathbb{R}$  a  $K_{t_0, x_0}^N$ -locally forced solution if it solves

$$(61) \quad \begin{cases} -\partial_{tt}w + \Delta w = (1_{K_{t_0, x_0}^N} F + w)^3 & (t, x) \in \mathbb{R} \times \mathbb{R}^4 \\ w|_{t=t_0} = w_0 \in \dot{H}^1(\mathbb{R}^4), \quad \partial_t w|_{t=t_0} = w_1 \in L^2(\mathbb{R}^4) \end{cases} .$$

We also require that the functions  $(w_0, w_1)$  agree with  $(v(t_0), \partial_t v(t_0))$  on the cube  $\|x - x_0\|_\infty \leq 2N$ .

**Remark 5.6.**

From finite speed of propagation, it follows that  $w|_{K_{t_0, x_0}^N} = v|_{K_{t_0, x_0}^N}$ .

For the same reasons as described in the last paragraph, we also use the energy on a slightly larger region. To this end, we define

$$\tilde{E}_{t_0, x_0}^N[w](t) := \int_{\|x - x_0\|_\infty \leq 16N - |t - t_0|} T^{00}[w](t, x) dx, \quad \text{where } t \in [t_0, t_0 + N] .$$

Thus, this definition is adapted to the fattened cone  $\tilde{K}_{t_0, x_0}^N$ , which is defined in (43). We also set

$$(62) \quad \tilde{\mathcal{E}}_{t_0, x_0}^N[w] := \sup_{t \in [t_0, t_0 + N]} \tilde{E}_{t_0, x_0}^N[w](t) .$$

**Lemma 5.7** (Local flux estimate).

Let  $t_0 \in N\mathbb{N}_0$ , let  $x_0 \in N\mathbb{Z}^4$ , and let  $w$  be a  $K_{t_0, x_0}^N$ -locally forced solution. Then, we have that

$$(63) \quad \sup_{\substack{t' \in [t_0, t_0 + N] \\ \|x' - x_0\|_{\ell^\infty} \leq 8N}} \int_{\substack{\|x - x'\|_2 = |t - t'| \\ t \in [t_0, t_0 + N]}} \frac{w^4}{4} d\sigma(t, x) \leq 4\tilde{\mathcal{E}}_{t_0, x_0}^N + 12 \int_{K_{t_0, x_0}^N} |F|(|w| + |F|)^2 |\partial_t w| dx dt .$$

We emphasize that, even though the energy  $\tilde{\mathcal{E}}_{t_0, x_0}^N$  is measured on a truncated  $\ell^\infty$ -cone, the flux is still controlled on a light cone. The estimate (63), however, only controls  $w$  on a lower-dimensional surface in space-time, and thus cannot directly be used to bound the energy increment. In our main argument, we rely on the following averaged version.

**Lemma 5.8** (Averaged local flux estimate).

Let  $N \geq N_{\text{hi}}$ , let  $t_0 \in N\mathbb{N}_0$ , let  $x_0 \in N\mathbb{Z}^4$ , and let  $w$  be a  $K_{t_0, x_0}^N$ -locally forced solution. Then, we have that

$$(64) \quad \begin{aligned} \tilde{\mathcal{F}}_{t_0, x_0}^N &:= \sup_{\substack{t' \in [t_0, t_0 + N] \\ \|x' - x_0\|_\infty \leq 4N}} \int_{\substack{\|x - x'\|_2 - |t - t'| \leq N^{10\delta} \\ t \in [t_0, t_0 + N]}} \frac{w^4}{4} \, dx dt \\ &\lesssim N^{40\delta} \left( \tilde{\mathcal{E}}_{t_0, x_0}^N + \int_{K_{t_0, x_0}^N} |F|(|w| + |F|)^2 |\partial_t w| \, dx dt \right). \end{aligned}$$

The appearance of  $N^{10\delta}$  is for technical reasons only, and the reader is encouraged to mentally replace it by 1. This will later help us to deal with the spatial tails of the wave packets.

*Proof of Lemma 5.7.* For the duration of this proof, we define

$$e(t) := \int_{\|x - x'\|_2 \leq |t - t'|} T^{00}(t, x) \, dx$$

From finite speed of propagation, we expect  $e(t)$  to be (nearly) non-increasing on  $[t_0, t']$  and non-decreasing on  $[t', t_0 + N]$ . From the assumptions above, it follows that  $\|x - x'\|_2 \leq |t - t'|$  implies

$$\|x - x_0\|_\infty \leq \|x - x'\|_\infty + \|x' - x_0\|_\infty \leq \|x - x'\|_2 + 8N \leq |t - t'| + 8N \leq 10N - |t - t_0|.$$

Thus, we obtain that  $e(t) \leq \tilde{\mathcal{E}}_{t_0, x_0}^N$  for all  $t \in [t_0, t_0 + N]$ . Using (54), we obtain for all  $t \in [t', t_0 + N]$  that

$$\begin{aligned} \frac{d}{dt} e(t) &= \int_{\|x - x'\|_2 \leq |t - t'|} \partial_t T^{00}(t, x) \, dx + \int_{\|x - x'\|_2 = |t - t'|} T^{00}(t, x) \, d\sigma(x) \\ &= - \int_{\|x - x'\|_2 \leq |t - t'|} \mathcal{N} \partial_t v \, dx - \int_{\|x - x'\|_2 \leq |t - t'|} \partial_{x_k} T^{0k} \, dx + \int_{\|x - x'\|_2 = |t - t'|} T^{00}(t, x) \, d\sigma(x) \\ &= - \int_{\|x - x'\|_2 \leq |t - t'|} \mathcal{N} \partial_t v \, dx + \int_{\|x - x'\|_2 = |t - t'|} (T^{00}(t, x) + T^{0k} \nu_k) \, d\sigma(x) \\ &\geq -6 \int_{\|x - x'\|_2 \leq |t - t'|} |F|(|F| + |v|)^2 |\partial_t v| \, dx + \int_{\|x - x'\|_2 = |t - t'|} \frac{v^4}{4} \, d\sigma(x). \end{aligned}$$

Integrating this inequality in time, we obtain the result on  $[t', t_0 + N]$ . The bound on  $[t_0, t']$  is similar.  $\square$

*Proof of Lemma 5.8.* Since  $N \geq N_{\text{hi}} \geq \eta^{-1}$ , we have that  $N^{10\delta} \ll N$ . We then simply integrate (63) over a spatial ball of radius  $\sim N^{10\delta}$  around  $x'$ .  $\square$

## 6. THE ENERGY INCREMENT AND INDUCTION ON SCALES

We are now ready to finally bound the energy increment of the nonlinear component  $v$ . The argument roughly splits into two parts: A single scale analysis and induction on scales.



For technical reasons, we define a flux-term involving a thinner neighborhood of the cone. More precisely, we let

$$\mathcal{F}_{t_0, x_0}^N[w] := \sup_{\substack{(t', x') \in [t_0, t_0 + N] \times \mathbb{R}^4 \\ \|x' - x_0\|_\infty \leq 4N}} \int_{\substack{\|x - x'\|_2 - |t - t'| \leq N^{5\delta} \\ t \in [t_0, t_0 + N]}} \frac{w^4}{4} dx dt$$

Recall that the light-cone in  $\tilde{\mathcal{F}}_{t_0, x_0}^N[w]$ , as defined in (64), has width  $N^{10\delta}$ .

**Proposition 6.1** (Single-scale energy increment).

Let  $N \geq 1$ . Let  $t_0 \in N\mathbb{N}_0$ , where  $0 \leq t_0/N \leq \lfloor N^\theta \rfloor$ , and let  $x_0 \in N\mathbb{Z}^4$ . Furthermore, let  $w_1, w_2 \in L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)$  and  $w_3 \in L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)$ . Then,

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} |F_M| |w_1| |w_2| |w_3| dx dt \\ & \lesssim \eta N^{\frac{3}{4} - s + 8\delta} (\|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{t_0, x_0}^N[w_1])^{\frac{1}{4}} (\|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{t_0, x_0}^N[w_2])^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)}. \end{aligned}$$

We have two separate reasons for introducing the auxiliary functions  $w_1$ ,  $w_2$ , and  $w_3$ . First, it emphasizes that the proof does not depend on the evolution equation of the nonlinear component. Second, it allows us to pass to smaller spatial scales than  $N$  with minimal notational effort, see Corollary 6.2.

*Proof.* If  $1 \leq N < N_{\text{hi}}$ , there is nothing to show. Thus, we may assume that  $N \geq N_{\text{hi}}$ .

*Step 1: Wave packet decomposition.* Recall from (35) and (36) that

$$F_N^\omega = \sum_k \epsilon_k \exp(it|\nabla|) f_k^+ + \sum_k \epsilon_k \exp(-it|\nabla|) f_k^-.$$

We only control the contribution of  $\sum_k \epsilon_k \exp(it|\nabla|) f_k^+$ , the other estimate is nearly identical. Then, we may also drop the superscript  $+$  from our notation. We now apply Proposition 4.1 and Proposition 4.3 to the family  $\{\epsilon_k f_k\}_k$ , and let the sets  $\mathcal{A}_m$ ,  $\mathcal{B}_{j,m}$ , and  $\mathcal{D}_m$  be as in Proposition 4.3. As before, we implicitly restrict to  $\|k\|_\infty \in (N/2, N]$ . We also write  $f_{k,l} = f_{k,l;t_0}$ .

*Step 2: Distant wave packets and extreme amplitudes.* On a heuristic level, the wave packets whose tubes  $T_{k,l}$  do not intersect  $K_{t_0, x_0}^N$  should not contribute to the integral. We now make this precise using the decay estimate (42). Indeed,

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} \left| \sum_{\substack{(k,l) \in \mathbb{Z}^4 \times \mathbb{Z}^4: \\ \|l - x_0\|_\infty > 3N}} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k,l} \right| |w_1| |w_2| |w_3| dx dt \\ & \lesssim N \left\| \sum_{\substack{(k,l) \in \mathbb{Z}^4 \times \mathbb{Z}^4: \\ \|l - x_0\|_\infty > 3N}} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k,l} \right\|_{L_t^\infty L_x^\infty(K_{t_0, x_0}^N)} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim N \left( \sum_{\substack{(k,l) \in \mathbb{Z}^4 \times \mathbb{Z}^4: \\ \|l - x_0\|_\infty > 3N}} N^{-100} (1 + \|x_0 - l\|_2)^{-100} \|f_k\|_{L_x^2(\mathbb{R}^4)} \right) \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim N^{-99+2} \left( \sum_{k \in \mathbb{Z}^4} \|f_k\|_{L_x^2(\mathbb{R}^4)}^2 \right)^{\frac{1}{2}} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \end{aligned}$$

$$\lesssim \eta N^{\frac{3}{4}-s} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} .$$

Thus, this contribution is acceptable. It remains to control the wave packets with indices in  $\bigcup_{m \in \mathbb{Z}} \mathcal{A}_m$ . We now use crude estimates to reduce to  $\sim \log(N)$  amplitude scales. Let  $m \leq -20 \log(N)$ . Since  $\#\mathcal{A}_m \lesssim N^8$ , we have that

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} \left| \sum_{(k, l) \in \mathcal{A}_m} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k, l} \right| |w_1| |w_2| |w_3| dx dt \\ & \lesssim N \left( \sum_{(k, l) \in \mathcal{A}_m} \|\exp(i(t - t_0)|\nabla|) f_{k, l}\|_{L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R}^4)} \right) \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim N 2^m \#\mathcal{A}_m \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim 2^m N^9 \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} . \end{aligned}$$

Summing this inequality over all  $m \leq -20 \log(N)$ , we obtain that

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} \left| \sum_{m \leq -20 \log(N)} \sum_{(k, l) \in \mathcal{A}_m} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k, l} \right| |w_1| |w_2| |w_3| dx dt \\ & \lesssim N^{-11} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim \eta N^{\frac{3}{4}-s} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} . \end{aligned}$$

Finally, if  $\mathcal{A}_m \neq \emptyset$ , then  $\#\mathcal{A}_m \geq 1$ . This implies that

$$2^m \leq 2^m (\#\mathcal{A}_m)^{\frac{1}{2}} \lesssim \left( \sum_{k, l} \|f_{k, l}\|_{L^2(\mathbb{R}^4)}^2 \right)^{\frac{1}{2}} \lesssim \eta N^{-s} .$$

For a sufficiently small absolute constant  $\eta > 0$ , this implies that  $m \leq 0$ . This completes the crude part of the argument. *Step 3: Bushes.* First, we define the fattened tubes by

$$\tilde{T}_{k, l} := \{(t, x) \in [t_0, t_0 + N] \times \mathbb{R}^d : |x - (l - t \cdot k / \|k\|_2)| \leq N^{2\delta}\} .$$

Furthermore, we define the collection of fattened tubes corresponding to a bush by

$$\tilde{T}(\mathcal{B}_{j, m}) := \bigcup_{(k, l) \in \mathcal{B}_{j, m}} \tilde{T}_{k, l} .$$

With these definitions in hand, we now write

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} \left| \sum_{j=1}^J \sum_{(k, l) \in \mathcal{B}_{j, m}} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k, l} \right| |w_1| |w_2| |w_3| dx dt \\ & \leq \sum_{j=1}^J \int_{\tilde{T}(\mathcal{B}_{j, m})} \left| \sum_{(k, l) \in \mathcal{B}_{j, m}} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k, l} \right| |w_1| |w_2| |w_3| dx dt \\ & + \sum_{j=1}^J \int_{K_{t_0, x_0}^N \setminus \tilde{T}(\mathcal{B}_{j, m})} \left| \sum_{(k, l) \in \mathcal{B}_{j, m}} \epsilon_k \exp(i(t - t_0)|\nabla|) f_{k, l} \right| |w_1| |w_2| |w_3| dx dt \end{aligned}$$

Using that all tubes in  $\mathcal{B}_{j,m}$  pass through the same space-time cube of size  $\sim N^\delta$ , we obtain from Proposition 4.8 that

$$\begin{aligned}
& \sum_{j=1}^J \int_{\tilde{T}(\mathcal{B}_{j,m})} \left| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right| |w_1| |w_2| |w_3| dx dt \\
& \lesssim \sum_{j=1}^J N^{\frac{1}{2}} \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right\|_{L_{t,x}^\infty(\mathbb{R} \times \mathbb{R}^4)} \|w_1\|_{L_{t,x}^4(\tilde{T}(\mathcal{B}_{j,m}))} \|w_2\|_{L_{t,x}^4(\tilde{T}(\mathcal{B}_{j,m}))} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim \eta N^{\frac{1}{2}+8\delta} 2^m \left( \sum_{j=1}^J (\#\mathcal{B}_{j,m})^{\frac{1}{2}} \right) \mathcal{F}_{t_0,x_0}^N [w_1]^{\frac{1}{4}} \mathcal{F}_{t_0,x_0}^N [w_2]^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim \eta N^{\frac{1}{2}+8\delta} 2^m \left( \sum_{j=1}^J \mathcal{B}_{j,m} \right) \mu^{-\frac{1}{2}} \mathcal{F}_{t_0,x_0}^N [w_1]^{\frac{1}{4}} \mathcal{F}_{t_0,x_0}^N [w_2]^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim \eta N^{\frac{1}{2}+8\delta} 2^m (\#\mathcal{A}_m) \mu^{-\frac{1}{2}} \mathcal{F}_{t_0,x_0}^N [w_1]^{\frac{1}{4}} \mathcal{F}_{t_0,x_0}^N [w_2]^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)}.
\end{aligned}$$

Here,  $\mu$  denotes the minimum number of packets inside a single bush, see Proposition 4.3. Using the decay estimate (42), we control the contributions outside the  $\tilde{T}(\mathcal{B}_{j,m})$  by

$$\begin{aligned}
& \sum_{j=1}^J \int_{K_{t_0,x_0}^N \setminus \tilde{T}(\mathcal{B}_{j,m})} \left| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right| |w_1| |w_2| |w_3| dx dt \\
& \lesssim N \sum_{j=1}^J \left\| \sum_{(k,l) \in \mathcal{B}_{j,m}} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right\|_{L_{t,x}^\infty(K_{t_0,x_0}^N \setminus \tilde{T}(\mathcal{B}_{j,m}))} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim N^1 N^{-100} 2^m \left( \sum_{j=1}^J \#\mathcal{B}_{j,m} \right) \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim \eta N^{\frac{3}{4}-s} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)}.
\end{aligned}$$

In the last line, we have used that  $2^m (\#\mathcal{A}_m)^{\frac{1}{2}} \lesssim \eta$  and  $\#\mathcal{A}_m \lesssim N^8$ .

*Step 4: Disjoint wave packets.* We now control the contribution of the almost disjoint family  $\mathcal{D}_m$ . If  $\mu < 1$ , then  $\mathcal{D}_m$  is empty, and there is nothing to prove. If  $\mu \geq 1$ , it follows from Proposition 4.8 that

$$\begin{aligned}
& \int_{K_{t_0,x_0}^N} \left| \sum_{(k,l) \in \mathcal{D}_m} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right| |w_1| |w_2| |w_3| dx dt \\
& \lesssim N \left\| \sum_{(k,l) \in \mathcal{D}_m} \epsilon_k \exp(i(t-t_0)|\nabla|) f_{k,l} \right\|_{L_{t,x}^\infty(K_{t_0,x_0}^N)} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\
& \lesssim \eta N^{1+8\delta} 2^m \mu^{\frac{1}{2}} \|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)}.
\end{aligned}$$

*Step 5: Finishing the proof.*

In total, we have shown that

$$(65) \quad \begin{aligned} & \int_{K_{t_0, x_0}^N} |F_M| |w_1| |w_2| |w_3| dx dt \\ & \lesssim \eta N^{8\delta} \left( N^{\frac{3}{4}-s} + N 2^m \mu^{\frac{1}{2}} + N^{\frac{1}{2}} 2^m (\#\mathcal{A}_m) \mu^{-\frac{1}{2}} \right) (\|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{t_0, x_0}^N[w_1])^{\frac{1}{4}} \\ & \quad \cdot (\|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{t_0, x_0}^N[w_2])^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)}. \end{aligned}$$

Due to our choice  $\mu := N^{-\frac{1}{2}} \#\mathcal{A}_m$  in Proposition 4.3, this completes the proof.  $\square$

**Corollary 6.2** (Coarse-scale energy increment).

Let  $N \geq 1$ . Let  $t_0 \in N\mathbb{N}_0$ , with  $0 \leq t_0/N \leq \lfloor N^\theta \rfloor$ , and let  $x_0 \in N\mathbb{Z}^4$ . Let  $w$  be a  $K_{t_0, x_0}^N$ -locally forced solution. Then, we have for all  $M \geq N$  that

$$(66) \quad \int_{K_{t_0, x_0}^N} |F_M| |w|^2 |\partial_t w| dx dt \lesssim \eta M^{\frac{3}{4}-s+8\delta} \tilde{\mathcal{E}}_{t_0, x_0}^N[w]^{\frac{1}{2}} \left( \tilde{\mathcal{E}}_{t_0, x_0}^N[w] + \tilde{\mathcal{F}}_{t_0, x_0}^N[w] \right)^{\frac{1}{2}}.$$

We refer to Corollary 6.2 as a coarse scale estimate since the wave packets in  $F_M$  are atleast as long as the length of  $K_{t_0, x_0}^N$  (in time).

*Proof.* As before, we may take  $M \geq N_{\text{hi}}$ . We distinguish two different cases. If  $M \geq N^{\frac{4}{3}}$ , we obtain from Proposition 4.8 that

$$\int_{K_{t_0, x_0}^N} |F_M| |w|^2 |\partial_t w| dx dt \leq N \|F_M\|_{L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R}^4)} \|w\|_{L_t^\infty L_x^4(K_{t_0, x_0}^N)}^2 \|\partial_t w\|_{L_t^\infty L_x^2(K_{t_0, x_0}^N)} \lesssim \eta M^{\frac{3}{4}-s+\delta} \tilde{\mathcal{E}}_{t_0, x_0}^N[w].$$

Next, we let  $M \leq N^{\frac{4}{3}}$ . Then, there exist  $\tau_0 \in M\mathbb{N}_0$  and  $y_0 \in M\mathbb{Z}^4$  s.t.  $K_{t_0, x_0}^N \subseteq K_{\tau_0, y_0}^M$ . Furthermore, since  $t_0/N \leq \lfloor N^\theta \rfloor$ , it holds that  $\tau_0/M \leq \lfloor M^\theta \rfloor$ . Set  $w_1 = w_2 := 1_{K_{t_0, x_0}^N} w$  and  $w_3 := 1_{K_{t_0, x_0}^N} w$ . Since  $M \leq N^{\frac{4}{3}}$ , we have that

$$\mathcal{F}_{\tau_0, y_0}^M[w_1] = \mathcal{F}_{\tau_0, y_0}^M[w_2] \leq \tilde{\mathcal{F}}_{t_0, x_0}^N[w]$$

Using the single-scale energy increment (Proposition 6.1), we obtain that

$$\begin{aligned} & \int_{K_{t_0, x_0}^N} |F_M| |w|^2 |\partial_t w| dx dt \\ & = \int_{K_{\tau_0, y_0}^M} |F_M| |w_1| |w_2| |w_3| dx dt \\ & \lesssim \eta M^{\frac{3}{4}-s+\delta} (\|w_1\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{\tau_0, y_0}^M[w_1])^{\frac{1}{4}} (\|w_2\|_{L_t^\infty L_x^4(\mathbb{R} \times \mathbb{R}^4)}^4 + \mathcal{F}_{\tau_0, y_0}^M[w_2])^{\frac{1}{4}} \|w_3\|_{L_t^\infty L_x^2(\mathbb{R} \times \mathbb{R}^4)} \\ & \lesssim \eta M^{\frac{3}{4}-s+\delta} \tilde{\mathcal{E}}_{t_0, x_0}^N[w]^{\frac{1}{2}} \left( \tilde{\mathcal{E}}_{t_0, x_0}^N[w] + \tilde{\mathcal{F}}_{t_0, x_0}^N[w] \right)^{\frac{1}{2}}. \end{aligned}$$

$\square$

Due to Proposition 6.1 and Corollary 6.2, we understand the energy increment at a single scale. Unfortunately, the cone  $K_{t_0, x_0}^N$  may contain many wave packets on smaller scales. Similar problems are often

encountered in restriction theory, and can sometimes be solved using Wolff's induction on scales strategy [57]. The following argument can be seen as a (simple) implementation of this idea.

**Proposition 6.3** (Induction on scales).

Let  $s > \max(1 - \theta/2, 3/4 + \theta)$ . Let  $R \geq 1$  be a dyadic integer, and let  $F$  be as in Proposition 4.8. Let  $t_0 \in \mathbb{R}\mathbb{N}_0$ , with  $t_0/R \leq \lfloor R^\theta \rfloor$ ,  $x_0 \in \mathbb{R}\mathbb{Z}^4$ , let  $w$  be a  $K_{t_0, x_0}^R$ -locally forced solution. For a large absolute constant  $C_1 \geq 1$ , we have that

$$(67) \quad \tilde{\mathcal{E}}_{t_0, x_0}^R[w] \leq 2E_{|x-x_0| \leq 16R}[w](t_0) + C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6$$

and

$$(68) \quad \tilde{\mathcal{F}}_{t_0, x_0}^R[w] \leq C_1 R^{50\delta} \left( E_{|x-x_0| \leq 16R}[w](t_0) + \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6 \right).$$

*Proof.* We use induction on the dyadic integers  $R \geq 1$ .

*Step 1: Base case  $R = 1$ .* We have that

$$\begin{aligned} \tilde{\mathcal{E}}_{t_0, x_0}^1[w] &\leq E_{|x-x_0| \leq 16}[w](t_0) + C \int_{K_{t_0, x_0}^1} |F||w|^2 |\partial_t w| dx dt + C \int_{K_{t_0, x_0}^1} |F|^3 |\partial_t w| dx dt \\ &\leq E_{|x-x_0| \leq 16}[w](t_0) + C \|F\|_{L_t^\infty L_x^\infty(\mathbb{R} \times \mathbb{R}^4)} \|w\|_{L_t^\infty L_x^4(K_{t_0, x_0}^1)}^2 \|\partial_t w\|_{L_t^\infty L_x^2(K_{t_0, x_0}^1)} \\ &\quad + C \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^1)}^3 \|\partial_t w\|_{L_t^\infty L_x^2(K_{t_0, x_0}^1)} \\ &\leq E_{|x-x_0| \leq 16}[w](t_0) + \frac{1}{4} \tilde{\mathcal{E}}_{t_0, x_0}^1[w] + C \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^1)}^3 \tilde{\mathcal{E}}_{t_0, x_0}^1[w]^{\frac{1}{2}} \\ &\leq E_{|x-x_0| \leq 16}[w](t_0) + \frac{1}{2} \tilde{\mathcal{E}}_{t_0, x_0}^1[w] + C \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^1)}^6. \end{aligned}$$

Insert this bound into Lemma 5.8, we obtain that

$$\begin{aligned} \tilde{\mathcal{F}}_{t_0, x_0}^1[w] &\lesssim \tilde{\mathcal{E}}_{t_0, x_0}^1[w] + \int_{K_{t_0, x_0}^1} |F||w|^2 |\partial_t w| dx dt + C \int_{K_{t_0, x_0}^1} |F|^3 |\partial_t w| dx dt \\ &\lesssim \tilde{\mathcal{E}}_{t_0, x_0}^1[w] \\ &\lesssim E_{|x-x_0| \leq 16}[w](t_0) + \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^1)}^6. \end{aligned}$$

By choosing  $C_1$  sufficiently large, we obtain (67) and (68). This already determines our choice of  $C_1$ , which we now regard as a fixed constant. Let  $R \geq 2$  be an arbitrary dyadic integer. Using the induction hypothesis, we can rely on the inequalities (67) and (68) for all scales  $N \leq R/2$ .

*Step 2: Splitting the energy increment.* From Lemma 5.4, we have that

$$(69) \quad \tilde{\mathcal{E}}_{t_0, x_0}^R[w] \leq E_{|x-x_0| \leq 16R}[w](t_0) + C \int_{K_{t_0, x_0}^R} |F||w|^2 |\partial_t w| dx dt + C \int_{K_{t_0, x_0}^R} |F|^3 |\partial_t w| dx dt.$$

The main term is the second summand in (69). We use a Littlewood-Paley type decomposition of the linear evolution and write

$$\begin{aligned} & \int_{K_{t_0, x_0}^R} |F||w|^2 |\partial_t w| dx dt \\ & \leq \sum_{N \geq R} \int_{K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt + \sum_{N \leq R/2} \int_{K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt . \end{aligned}$$

*Step 3: High frequencies.* The high frequencies can be controlled using the single-scale estimate from Proposition 6.1. Indeed, we have that

$$\begin{aligned} (70) \quad & \sum_{N \geq R} \int_{K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt \\ & \lesssim \eta \sum_{N \geq R} N^{\frac{3}{4}-s+8\delta} \tilde{\mathcal{E}}_{t_0, x_0}^R[w]^{\frac{1}{2}} \left( \tilde{\mathcal{E}}_{t_0, x_0}^R[w] + \tilde{\mathcal{F}}_{t_0, x_0}^R[w] \right)^{\frac{1}{2}} \\ & \lesssim \eta R^{\frac{3}{4}-s+8\delta} \tilde{\mathcal{E}}_{t_0, x_0}^R[w]^{\frac{1}{2}} \left( \tilde{\mathcal{E}}_{t_0, x_0}^R[w] + \tilde{\mathcal{F}}_{t_0, x_0}^R[w] \right)^{\frac{1}{2}} . \end{aligned}$$

*Step 4: Low frequencies.* For  $\tau_n = Nn \in N\mathbb{N}_0$  and  $y_j = Nj \in N\mathbb{Z}^4$ , we write

$$\begin{aligned} & \int_{K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt \\ & = \sum_{n=0}^{\lfloor N^\theta \rfloor} \left( \int_{([\tau_n, \tau_n + N] \times \mathbb{R}^4) \cap K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt \right) + \int_{((N^{1+\theta}, \infty) \times \mathbb{R}^4) \cap K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt \\ & = \sum_{n=0}^{\lfloor N^\theta \rfloor} \sum_{\substack{j \in \mathbb{Z}^4 \\ K_{\tau_n, y_j}^N \subseteq K_{t_0, x_0}^R}} \int_{K_{\tau_n, y_j}^N} |F_N||w|^2 |\partial_t w| dx dt + \int_{((N^{1+\theta}, \infty) \times \mathbb{R}^4) \cap K_{t_0, x_0}^R} |F_N||w|^2 |\partial_t w| dx dt . \end{aligned}$$

In the last line, we have used that

$$K_{t_0, x_0}^R = \bigcup_{\substack{(n, j) \in \mathbb{N}_0 \times \mathbb{Z}^4 \\ K_{\tau_n, y_j}^N \subseteq K_{t_0, x_0}^R}} K_{\tau_n, y_j}^N .$$

We first control the contributions on the time intervals  $[\tau_n, \tau_n + N]$ . To this end, we define  $w^{(N, n, j)}$  as the  $K_{\tau_n, y_j}^N$ -locally forced solution with data

$$w^{(N, n, j)}(\tau_n) = w(\tau_n) \quad \text{and} \quad \partial_t w^{(N, n, j)}(\tau_n) = \partial_t w(\tau_n) .$$

Using finite speed of propagation,  $w$  and  $w^{(N, n, j)}$  coincide on  $K_{\tau_n, y_j}^N$ . Applying Proposition 6.1 and the induction hypothesis to  $w^{(N, n, j)}$ , it follows that

$$\begin{aligned}
& \int_{K_{\tau_n, y_j}^N} |F_N| |w|^2 |\partial_t w| dx dt \\
&= \int_{K_{\tau_n, y_j}^N} |F_N| |w^{(N, n, j)}|^2 |\partial_t w^{(N, n, j)}| dx dt \\
&\lesssim \eta N^{\frac{3}{4}-s+8\delta} \tilde{\mathcal{E}}_{\tau_n, y_j}^N [w^{(N, n, j)}]^{\frac{1}{2}} \left( \tilde{\mathcal{E}}_{\tau_n, y_j}^N [w^{(N, n, j)}]^{\frac{1}{2}} + \tilde{\mathcal{F}}_{\tau_n, y_j}^N [w^{(N, n, j)}] \right)^{\frac{1}{2}} \\
&\lesssim \eta N^{\frac{3}{4}-s+8\delta} \left( 2E_{|x-y_j| \leq 16N} [w^{(N, n, j)}](\tau_n) + C_1 \|F\|_{L_t^3 L_x^6(K_{\tau_n, y_j}^N)}^6 \right)^{\frac{1}{2}} \\
&\quad \cdot N^{50\delta} \left( C_1 E_{|x-y_j| \leq 16N} [w^{(N, n, j)}](\tau_n) + C_1 \|F\|_{L_t^3 L_x^6(K_{\tau_n, y_j}^N)}^6 \right)^{\frac{1}{2}} \\
&\lesssim \eta N^{\frac{3}{4}-s+60\delta} C_1 \left( E_{|x-y_j| \leq 16N} [w^{(N, n, j)}](\tau_n) + \|F\|_{L_t^3 L_x^6(K_{\tau_n, y_j}^N)}^6 \right) \\
&\lesssim \eta N^{\frac{3}{4}-s+60\delta} C_1 \left( E_{|x-y_j| \leq 16N} [w](\tau_n) + \|F\|_{L_t^3 L_x^6(K_{\tau_n, y_j}^N)}^6 \right).
\end{aligned}$$

As a consequence, we obtain that

$$\begin{aligned}
& \sum_{n=0}^{\lfloor N^\theta \rfloor} \sum_{\substack{j \in \mathbb{Z}^4 \\ K_{\tau_n, y_j}^N \subseteq K_{t_0, x_0}^R}} \int_{K_{\tau_n, y_j}^N} |F_N| |w|^2 |\partial_t w| dx dt \\
&\lesssim \eta N^{\frac{3}{4}-s+60\delta} C_1 \sum_{n=0}^{\lfloor N^\theta \rfloor} \sum_{\substack{j \in \mathbb{Z}^4 \\ K_{\tau_n, y_j}^N \subseteq K_{t_0, x_0}^R}} \left( E_{|x-y_j| \leq 16N} [w](\tau_n) + \|F\|_{L_t^3 L_x^6(K_{\tau_n, y_j}^N)}^6 \right) \\
&\lesssim \eta N^{\frac{3}{4}-s+60\delta} C_1 \sum_{n=\max(0, t_0/N)}^{\lfloor N^\theta \rfloor} \left( E_{|x-x_0| \leq 16R-|t-t_0|} [w](\tau_n) + \|F\|_{L_t^3 L_x^6(([\tau_n, \tau_n+N] \times \mathbb{R}^4) \cap K_{t_0, x_0}^R)}^6 \right) \\
(71) \quad &\lesssim \eta N^{\frac{3}{4}+\theta-s+60\delta} C_1 \tilde{\mathcal{E}}_{t_0, x_0}^R + \eta N^{\frac{3}{4}-s+60\delta} C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6.
\end{aligned}$$

Using the long-time decay estimate, we can control the contribution on the interval  $[N^{1+\theta}, \infty)$  by

$$\begin{aligned}
& \int_{([N^{1+\theta}, \infty) \times \mathbb{R}^4] \cap K_{t_0, x_0}^R} |F_N| |w|^2 |\partial_t w| dx dt \\
&\lesssim \|F_N\|_{L_t^1 L_x^\infty([N^{1+\theta}, \infty) \times \mathbb{R}^4)} \|w\|_{L_t^\infty L_x^4(K_{t_0, x_0}^R)}^2 \|\partial_t w\|_{L_t^\infty L_x^2(K_{t_0, x_0}^R)} \\
(72) \quad &\lesssim \eta N^{1-\frac{\theta}{2}-s+\delta} \tilde{\mathcal{E}}_{t_0, x_0}^R.
\end{aligned}$$

Combining (71) and (72), it follows that

$$\begin{aligned}
(73) \quad & \sum_{N \leq R/2} \int_{K_{t_0, x_0}^R} |F_N| |w|^2 |\partial_t w| dx dt \\
& \lesssim \eta C_1 \sum_{N \leq R/2} \left( N^{\frac{3}{4} + \theta - s + 5\delta} + N^{1 - \frac{\theta}{2} - s} \right) \tilde{\mathcal{E}}_{t_0, x_0}^R + \eta \left( \sum_{N \leq R/2} N^{\frac{3}{4} - s + 5\delta} \right) C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6 \\
& \lesssim \eta C_1 \tilde{\mathcal{E}}_{t_0, x_0}^R + \eta C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6.
\end{aligned}$$

Here, we have used that  $s > \max(1 - \frac{\theta}{2}, \frac{3}{4} + \theta)$ , and that  $\delta = \delta(s, \theta) > 0$  is sufficiently small.

*Step 5: Finishing the proof.* At this point, we have proven all the necessary estimates on  $w$ . It only remains to put them together, and use a “kick back” argument. From the energy increment (69), the high frequency estimate (70), the low-frequency estimate (73), and Hölder’s inequality, it follows that

$$(74) \quad \tilde{\mathcal{E}}_{t_0, x_0}^R[w] \leq E_{|x-x_0| \leq 16R}[w](t_0) + \eta C_1 \tilde{\mathcal{E}}_{t_0, x_0}^R[w] + \eta R^{\frac{3}{4} - s + 8\delta} \tilde{\mathcal{F}}_{t_0, x_0}^R[w] + \eta C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6$$

Inserting the same bound for the energy increment into (5.8), we also have that

$$(75) \quad \tilde{\mathcal{F}}_{t_0, x_0}^R[w] \lesssim R^{40\delta} \left( \tilde{\mathcal{E}}_{t_0, x_0}^R[w] + \eta R^{\frac{3}{4} - s + 8\delta} \tilde{\mathcal{F}}_{t_0, x_0}^R[w] + \eta C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6 \right)$$

If the absolute constant  $\eta = \eta(C_1, \delta) > 0$  is chosen sufficiently small, then (75) implies that

$$(76) \quad \tilde{\mathcal{F}}_{t_0, x_0}^R[w] \lesssim R^{40\delta} \left( \tilde{\mathcal{E}}_{t_0, x_0}^R[w] + \eta C_1 \|F\|_{L_t^3 L_x^6(K_{t_0, x_0}^R)}^6 \right).$$

Inserting this into (74), we obtain (67). Finally, (67) and (76) imply (68). This completes the proof of the induction step.  $\square$

Using Proposition 6.3, we now provide a short proof of the main result.

*Proof of Theorem 1.3.* Assume that the statements in Proposition 4.8 hold for  $\omega \in \Omega$ . From Lemma 5.1, it follows that there exists a local solution to (53). From Proposition 6.3, it follows for all  $R \geq 1$  that

$$\sup_{t \in [0, R]} \int_{|x| \leq 2R-t} T^{00}[v](t, x) dx \leq 2E[v_0, v_1] + C_1.$$

By letting  $R \rightarrow \infty$ , we obtain the a-priori energy bound

$$\sup_{t \in [0, \infty)} E[v](t) \leq 2E[v_0, v_1] + C_1.$$

From Proposition 5.2, this implies the global space-time bound  $\|v\|_{L_t^3 L_x^6([0, \infty) \times \mathbb{R}^4)} < \infty$  and the existence of scattering states  $(v_0^+, v_1^+) \in \dot{H}^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4)$ . Since  $u = F + v$ , we obtain the global space-time bound  $\|u\|_{L_t^3 L_x^6([0, \infty) \times \mathbb{R}^4)} < \infty$  and the scattering states  $(u_0^+, u_1^+) = (v_0^+ - f_{0, < N_{\text{hi}}}^\omega, v_1^+ - f_{1, < N_{\text{hi}}}^\omega)$ . This completes the proof for positive times. By time-reflection symmetry, we obtain the same result for negative times.  $\square$



## REFERENCES

- [1] Hajer Bahouri and Patrick Gérard. High frequency approximation of solutions to critical nonlinear wave equations. *Amer. J. Math.*, 121(1):131–175, 1999.
- [2] Árpád Bényi, Tadahiro Oh, and Oana Pocovnicu. On the probabilistic Cauchy theory of the cubic nonlinear Schrödinger equation on  $\mathbb{R}^d$ ,  $d \geq 3$ . *Trans. Amer. Math. Soc. Ser. B*, 2:1–50, 2015.
- [3] Árpád Bényi, Tadahiro Oh, and Oana Pocovnicu. Wiener randomization on unbounded domains and an application to almost sure well-posedness of NLS. In *Excursions in harmonic analysis. Vol. 4*, Appl. Numer. Harmon. Anal., pages 3–25. Birkhäuser/Springer, Cham, 2015.
- [4] Árpád Bényi, Tadahiro Oh, and Oana 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*, 6:114–160, 2019.
- [5] Árpád Bényi, Tadahiro Oh, and Oana Pocovnicu. On the probabilistic Cauchy theory for nonlinear dispersive PDEs. In *Landscapes of time-frequency analysis*, Appl. Numer. Harmon. Anal., pages 1–32. Birkhäuser/Springer, Cham, 2019.
- [6] Jean Bourgain. Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.*, 1(2):147–187, 1991.
- [7] Jean Bourgain. Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys.*, 166(1):1–26, 1994.
- [8] Jean Bourgain. Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. *Comm. Math. Phys.*, 176(2):421–445, 1996.
- [9] Jean Bourgain. Global wellposedness of defocusing critical nonlinear Schrödinger equation in the radial case. *J. Amer. Math. Soc.*, 12(1):145–171, 1999.
- [10] Jean Bourgain and Aynur Bulut. Invariant Gibbs measure evolution for the radial nonlinear wave equation on the 3d ball. *J. Funct. Anal.*, 266(4):2319–2340, 2014.
- [11] Bjoern Bringmann. Almost sure local well-posedness for a derivative nonlinear wave equation, September 2018.
- [12] Bjoern Bringmann. Almost-sure scattering for the radial energy-critical nonlinear wave equation in three dimensions. *Anal. PDE*, 13(4):1011–1050, 2020.
- [13] Nicolas Burq and Nikolay Tzvetkov. Random data Cauchy theory for supercritical wave equations. I. Local theory. *Invent. Math.*, 173(3):449–475, 2008.
- [14] Nicolas Burq and Nikolay Tzvetkov. Random data Cauchy theory for supercritical wave equations. II. A global existence result. *Invent. Math.*, 173(3):477–496, 2008.
- [15] Sagun Chanillo, Magdalena Czubak, Dana Mendelson, Andrea Nahmod, and Gigliola Staffilani. Almost sure boundedness of iterates for derivative nonlinear wave equations, October 2017.
- [16] Michael Christ, James Colliander, and Terence Tao. Ill-posedness for nonlinear Schrödinger and wave equations, November 2003.
- [17] James Colliander, Markus Keel, Gigliola Staffilani, Hideo Takaoka, and Terence Tao. Global well-posedness and scattering for the energy-critical nonlinear Schrödinger equation in  $\mathbb{R}^3$ . *Ann. of Math. (2)*, 167(3):767–865, 2008.
- [18] Antonio Cordoba. The Kakeya maximal function and the spherical summation multipliers. *Amer. J. Math.*, 99(1):1–22, 1977.
- [19] Giuseppe Da Prato and Arnaud Debussche. Two-dimensional Navier-Stokes equations driven by a space-time white noise. *J. Funct. Anal.*, 196(1):180–210, 2002.
- [20] Benjamin Dodson. Global well-posedness and scattering for the defocusing,  $L^2$ -critical nonlinear Schrödinger equation when  $d \geq 3$ . *J. Amer. Math. Soc.*, 25(2):429–463, 2012.
- [21] Benjamin Dodson. Global well-posedness and scattering for the defocusing,  $L^2$  critical, nonlinear Schrödinger equation when  $d = 1$ . *Amer. J. Math.*, 138(2):531–569, 2016.

- [22] Benjamin Dodson. Global well-posedness and scattering for the defocusing,  $L^2$ -critical, nonlinear Schrödinger equation when  $d = 2$ . *Duke Math. J.*, 165(18):3435–3516, 2016.
- [23] Benjamin Dodson. Global well-posedness and scattering for the defocusing, mass-critical generalized KdV equation. *Ann. PDE*, 3(1):Art. 5, 35, 2017.
- [24] Benjamin Dodson. Global well-posedness and scattering for the radial, defocusing, cubic nonlinear wave equation, September 2018.
- [25] Benjamin Dodson, Jonas Lührmann, and Dana Mendelson. Almost sure local well-posedness and scattering for the 4D cubic nonlinear Schrödinger equation. *Adv. Math.*, 347:619–676, 2019.
- [26] Benjamin Dodson, Jonas Lührmann, and Dana Mendelson. Almost sure scattering for the 4D energy-critical defocusing nonlinear wave equation with radial data. *Amer. J. Math.*, 142(2):475–504, 2020.
- [27] Charles Fefferman. A note on spherical summation multipliers. *Israel J. Math.*, 15:44–52, 1973.
- [28] Manoussos G. Grillakis. Regularity and asymptotic behaviour of the wave equation with a critical nonlinearity. *Ann. of Math. (2)*, 132(3):485–509, 1990.
- [29] Manoussos G. Grillakis. Regularity for the wave equation with a critical nonlinearity. *Comm. Pure Appl. Math.*, 45(6):749–774, 1992.
- [30] Larry Guth. A restriction estimate using polynomial partitioning. *J. Amer. Math. Soc.*, 29(2):371–413, 2016.
- [31] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.
- [32] Rowan Killip, Jason Murphy, and Monica Visan. Almost sure scattering for the energy-critical NLS with radial data below  $H^1(\mathbb{R}^4)$ . *Comm. Partial Differential Equations*, 44(1):51–71, 2019.
- [33] Rowan Killip, Terence Tao, and Monica Visan. The cubic nonlinear Schrödinger equation in two dimensions with radial data. *J. Eur. Math. Soc. (JEMS)*, 11(6):1203–1258, 2009.
- [34] Rowan Killip, Monica Visan, and Xiaoyi Zhang. The mass-critical nonlinear Schrödinger equation with radial data in dimensions three and higher. *Anal. PDE*, 1(2):229–266, 2008.
- [35] Sergiu Klainerman and Daniel Tataru. On the optimal local regularity for Yang-Mills equations in  $\mathbf{R}^{4+1}$ . *J. Amer. Math. Soc.*, 12(1):93–116, 1999.
- [36] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.
- [37] Jonas Lührmann and Dana Mendelson. Random data Cauchy theory for nonlinear wave equations of power-type on  $\mathbb{R}^3$ . *Comm. Partial Differential Equations*, 39(12):2262–2283, 2014.
- [38] Jonas Lührmann and Dana Mendelson. On the almost sure global well-posedness of energy sub-critical nonlinear wave equations on  $\mathbb{R}^3$ . *New York J. Math.*, 22:209–227, 2016.
- [39] Jason Murphy. Random data final-state problem for the mass-subcritical NLS in  $L^2$ . *Proc. Amer. Math. Soc.*, 147(1):339–350, 2019.
- [40] Andrea R. Nahmod, Tadahiro Oh, Luc Rey-Bellet, and Gigliola Staffilani. 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.
- [41] Andrea R. Nahmod, Nataša Pavlović, and Gigliola Staffilani. Almost sure existence of global weak solutions for supercritical Navier-Stokes equations. *SIAM J. Math. Anal.*, 45(6):3431–3452, 2013.
- [42] Tadahiro Oh, Mamoru Okamoto, and Oana Pocovnicu. On the probabilistic well-posedness of the nonlinear Schrödinger equations with non-algebraic nonlinearities. *Discrete Contin. Dyn. Syst.*, 39(6):3479–3520, 2019.
- [43] Tadahiro Oh and Oana Pocovnicu. Probabilistic global well-posedness of the energy-critical defocusing quintic nonlinear wave equation on  $\mathbb{R}^3$ . *J. Math. Pures Appl. (9)*, 105(3):342–366, 2016.
- [44] Oana 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(8):2521–2575, 2017.
- [45] Jeffrey Rauch. I. The  $u^5$  Klein-Gordon equation. II. Anomalous singularities for semilinear wave equations. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. I (Paris,*

- 1978/1979), volume 53 of *Res. Notes in Math.*, pages 335–364. Pitman, Boston, Mass.-London, 1981.
- [46] Eric Ryckman and Monica Visan. Global well-posedness and scattering for the defocusing energy-critical nonlinear Schrödinger equation in  $\mathbb{R}^{1+4}$ . *Amer. J. Math.*, 129(1):1–60, 2007.
  - [47] Jalal Shatah and Michael Struwe. Regularity results for nonlinear wave equations. *Ann. of Math. (2)*, 138(3):503–518, 1993.
  - [48] Jalal Shatah and Michael Struwe. Well-posedness in the energy space for semilinear wave equations with critical growth. *Internat. Math. Res. Notices*, (7):303ff., approx. 7 pp. 1994.
  - [49] Walter A. Strauss. Decay and asymptotics for  $cmu = F(u)$ . *J. Functional Analysis*, 2:409–457, 1968.
  - [50] Michael Struwe. Globally regular solutions to the  $u^5$  Klein-Gordon equation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(3):495–513 (1989), 1988.
  - [51] Terence Tao. Some recent progress on the restriction conjecture. In *Fourier analysis and convexity*, Appl. Numer. Harmon. Anal., pages 217–243. Birkhäuser Boston, Boston, MA, 2004.
  - [52] Terence Tao. Spacetime bounds for the energy-critical nonlinear wave equation in three spatial dimensions. *Dyn. Partial Differ. Equ.*, 3(2):93–110, 2006.
  - [53] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed sensing*, pages 210–268. Cambridge Univ. Press, Cambridge, 2012.
  - [54] Roman Vershynin. *High-dimensional probability*, volume 47 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2018. An introduction with applications in data science, With a foreword by Sara van de Geer.
  - [55] Monica Visan. The defocusing energy-critical nonlinear Schrödinger equation in higher dimensions. *Duke Math. J.*, 138(2):281–374, 2007.
  - [56] Thomas Wolff. Recent work connected with the Kakeya problem. In *Prospects in mathematics (Princeton, NJ, 1996)*, pages 129–162. Amer. Math. Soc., Providence, RI, 1999.
  - [57] Thomas Wolff. A sharp bilinear cone restriction estimate. *Ann. of Math. (2)*, 153(3):661–698, 2001.

Bjoern Bringmann, UNIVERSITY OF CALIFORNIA, LOS ANGELES, DEPARTMENT OF MATHEMATICS, 520 PORTOLA PLAZA, LOS ANGELES, CA 90095  
Email address: [bringmann@math.ucla.edu](mailto:bringmann@math.ucla.edu)