J. Eur. Math. Soc. (Online first) DOI 10.4171/JEMS/1314

Published by EMS Press

Chenjie Fan · Dana Mendelson

Construction of L^2 log-log blowup solutions for the mass critical nonlinear Schrödinger equation

Received November 2, 2020; revised April 27, 2021

Abstract. In this article, we study the log-log blowup dynamics for the mass critical nonlinear Schrödinger equation on $\mathbb{R}^2$ under rough but structured random perturbations at $L^2(\mathbb{R}^2)$ regularity. In particular, by employing probabilistic methods, we provide a construction of a family of $L^2(\mathbb{R}^2)$ regularity solutions which do not lie in $H^s(\mathbb{R}^2)$ for any $s > 0$, and which blowup according to the log-log dynamics.

Keywords. NLS, log-log blowup, random data

1. Introduction

1.1. Main results and background

We consider the focusing cubic nonlinear Schrödinger equation (NLS) on $\mathbb{R}^2$

$$\begin{cases} iu_t + \Delta u = -|u|^2 u, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^2, \\ u(0,x) = u_0. \end{cases} \tag{1.1}$$

The goal of this article is to construct log-log blowup solutions at $L^2_x(\mathbb{R}^2)$ regularity via random data methods.

The NLS (1.1) has three conservation laws:

- Mass: $M(u) = \int |u|^2$,
- Momentum: $P(u) = \Im \int u \nabla \bar{u}$,
- Energy: $E(u) = \frac{1}{2} \int |\nabla u|^2 - \frac{1}{4} \int |u|^4$,

Chenjie Fan: Academy of Mathematics and Systems Science, Beijing, 100080 China; fancj@amss.ac.cn

Dana Mendelson: Department of Mathematics, University of Chicago, Chicago, IL 60637, United States; dana@math.uchicago.edu

Mathematics Subject Classification (2020): Primary 35B44; Secondary 35Q55

and enjoys the scaling, translation and phase symmetries. In particular, if u solves (1.1) with initial data u_0, then

$$\frac{1}{\lambda_0}u\left(\frac{t}{\lambda_0^2}, \frac{x-x_0}{\lambda_0}\right)e^{i\gamma_0}, \quad x_0 \in \mathbb{R}^2,\ \lambda_0 > 0,\ \gamma_0 \in \mathbb{R},$$

solves (1.1) with initial data

$$u_{0,\lambda_0} = \frac{1}{\lambda_0}u_0\left(\frac{x-x_0}{\lambda_0}\right)e^{i\gamma_0}.$$

One may verify that the mass $M(u)$ is invariant under the same scaling symmetry, and hence equation (1.1) is referred to as the mass critical NLS. We note that solutions of (1.1) also enjoy the so-called Galilean and pseudoconformal symmetries. We will not explicitly use these symmetries in the present work, even though we will rely on many previous results on log-log blowup solutions for which these symmetries play a central role.

It is classical that $L_x^2(\mathbb{R}^2)$ initial data gives rise to unique local-in-time solutions of (1.1); see [17]. The focusing nature of (1.1) implies, in particular, the existence of a ground state solution, $Q(x)$, which is the unique $L^2(\mathbb{R}^2)$ radial positive solution of

$$-\Delta Q + Q = Q^3. \tag{1.2}$$

The ground state plays an essential role in the blowup behavior of (1.1), and in particular it provides a threshold for blowup dynamics in the following sense: for all $L_x^2(\mathbb{R}^2)$ solutions with mass strictly below $\|Q\|_{L_x^2}$, the associated flow is global and scatters asymptotically; see the work of Weinstein [51] and Dodson [22]. Moreover, there exists an explicit blowup solution with mass equal to $\|Q\|_{L_x^2}^2$, given by

$$S(t,x) = \frac{1}{t}Q\left(\frac{x}{t}\right)e^{-i/t+i\frac{|x|^2}{4t}}, \tag{1.3}$$

which, in some sense, is the unique minimal mass blowup solution; see [35].

Classical virial identity arguments due to Glassey [25] establish the existence of a large family of negative energy blowup solutions, but the argument does not directly characterize the blowup mechanism for such solutions. It is an active area of research to understand blowup for (1.1) from a constructive perspective, so that one may better understand possible blowup mechanisms in general. For blowup solutions with mass slightly above the ground state,

$$\|Q\|_2 < \|u_0\|_2 < \|Q\|_2 + \alpha^*, \tag{1.4}$$

where α^* is a small universal number, one of the best understood blowup dynamics is the so-called *log-log blowup*. Log-log blowup solutions have been studied numerically in [31], and the first mathematical construction of such solutions was provided by Perelman [44]. These solutions were subsequently systematically studied by Merle and Raphaël [36–38, 40]. In particular, Merle and Raphaël prove that for all $H_x^1(\mathbb{R}^2)$ solutions to (1.1) which have nonpositive energy[1] and with mass slightly above ground state, i.e. in the range (1.4), such solutions will blow up in finite time[2] $T < \infty$, with precise asymptotics as t approaches the blowup time T given by the following:

Definition 1.1 (Log-log blowup dynamics).

$$u(t,x) = \frac{1}{\lambda(t)}(Q+\epsilon)\left(\frac{x-x(t)}{\lambda(t)}\right)e^{-i\gamma(t)}, \qquad \frac{1}{\lambda(t)} \sim \sqrt{\frac{\log|\log(T-t)|}{T-t}} \tag{1.5}$$

and where

$$\epsilon(t) \xrightarrow{t\to T} 0 \quad \text{in } \dot{H}^1(\mathbb{R}^2) \cap L^2_{\text{loc}}(\mathbb{R}^2).$$

Such blowup was shown to be stable in $H^1_x(\mathbb{R}^2)$ in [47], and was later proved to be stable under $H^s_x(\mathbb{R}^2)$ perturbations, for all $s > 0$, by Colliander and Raphaël [20], though one needs to reformulate the notion of log-log blowup (in a natural way) for infinite energy solutions. It is unclear whether such blowup is stable under $L^2_x(\mathbb{R}^2)$ perturbations, although one may guess that the answer is negative given the result of [39].

In light of the speculation that stability of log-log blowup may be false for arbitrary data in $L^2_x(\mathbb{R}^2)$, and the fact that this long-standing question remains open, in the current work we investigate the stability of log-log blowup solutions under *random* $L^2_x(\mathbb{R}^2)$ perturbations. Beginning with the seminal work of Bourgain [5], the behavior of nonlinear dispersive equations with random initial data has been an active field of research; see further discussion in Section 1.2 below. Indeed, in spite of the absence of known deterministic well-posedness theory, or even the existence of ill-posedness results, randomization often lets one establish that a given dispersive equation is well-posed *almost surely* in a particular low-regularity function space.

In the current article we employ randomization for a different and novel purpose. Our aim is to establish the existence of blowup solutions at $L^2_x(\mathbb{R}^2)$ regularity via a probabilistic construction. We state our main theorem non-technically for the time being:

Theorem 1.2. *The log-log blowup dynamics of Definition 1.1 is stable, with high probability, under* (*certain structured*) *random* $L^2_x(\mathbb{R}^2)$ *perturbations.*

We will begin with well-prepared $H^1_x(\mathbb{R}^2)$ data which are known to lead to log-log blowup, and we perturb this initial data with random initial data, constructed as follows: Let $\{g_k\}_{k\in\mathbb{Z}^2}$ be a sequence of iid complex Gaussian mean-zero random variables. Let $\{P_k\}_{k\in\mathbb{Z}^2}$ be unit-scale projections to frequency $k \in \mathbb{Z}^2$, defined as the Fourier multiplier with respect to translations of a *fixed* Schwartz function

$$\psi_k(\xi) := \psi(\xi - k), \tag{1.6}$$

[1] The result of Merle and Raphaël is more general and this nonpositive energy assumption can be relaxed. However, general positive energy solutions are less understood compared to those blowing up according to the log-log law.

[2] It is already highly nontrivial that such a solution will blow up in finite time.

that is,

$$P_k f = \mathcal{F}^{-1}(\psi_k(\xi)\widehat{f}(\xi)). \tag{1.7}$$

We crucially exploit the fact that these Fourier projections satisfy a unit-scale Bernstein inequality, namely for all $1 \leq r_1 \leq r_2 \leq \infty$ we have

$$\|P_k f\|_{L_x^{r_2}(\mathbb{R}^2)} \leq C(r_1, r_2)\|P_k f\|_{L_x^{r_1}(\mathbb{R}^2)} \tag{1.8}$$

with a constant which is independent of $k \in \mathbb{Z}^4$.

Let $f \in L_x^2(\mathbb{R}^2)$, and define its randomization

$$f^\omega = \sum_{k\in\mathbb{Z}^2} g_k(\omega) P_k f = \sum_{k\in\mathbb{Z}^2} g_k(\omega)(\psi_k \widehat{f})^\vee. \tag{1.9}$$

Similar randomizations have previously been used in Euclidean space, first in [52], and subsequently in [2,33]. One can show that if $f \in L_x^2(\mathbb{R}^2) \setminus H_x^s(\mathbb{R}^2)$ for some $s > 0$, then $f^\omega \in L_x^2(\mathbb{R}^2) \setminus H_x^s(\mathbb{R}^2)$ almost surely, and throughout, we will restrict to the subset of full measure of Ω so that this is indeed the case without further comment.

In the present application, we will take f to be piecewise constant in Fourier space, i.e. $f_k := P_k f$ constant, and we further require that f_k satisfy[3]

$$|f_k| \leq C/|k|, \quad k \geq 1. \tag{1.10}$$

Additionally, we normalize

$$\sum_k |f_k|^2 = 1. \tag{1.11}$$

Note that in particular there are many $L_x^2(\mathbb{R}^2)$ functions f with this property which do not belong to $H^s(\mathbb{R}^2)$ for any $s > 0$, and hence f^ω does not belong to $H_x^s(\mathbb{R}^2)$ for any $s > 0$: consider for instance the function f which satisfies

$$|f_k| \sim \frac{1}{|k|\log^2|k|}, \quad k \in \mathbb{Z}^2.$$

We note that our result works almost line by line if one assumes f_k is a function rather than a number. Then one needs to replace $|f_k|$ in (1.10) and (1.11) by $\|f_k\|_{L_x^\infty}$.

We will provide more details about the precise form of the $H_x^1(\mathbb{R}^2)$ blowup data in Section 3, and we will state a more detailed version of Theorem 1.2 in Theorem 3.1 below.

Remark 1.3. While our techniques are probabilistic, in light of the previous discussion on the randomized initial data, our main theorem provides a construction of $L_x^2(\mathbb{R}^2)$ log-log blowup solutions for the mass critical nonlinear Schrödinger equation which do not lie in $H_x^s(\mathbb{R}^2)$ for any $s > 0$. To the best our knowledge, such examples were not previously known in the literature. Consequently, one can view our result as an example of the probabilistic method, whose use in combinatorics was pioneered by Erdős [24].

[3] This is to mimic the randomization in [5].

Remark 1.4. We emphasize that unlike many random data results, (1.1) is locally well-posed at $L^2_x(\mathbb{R}^2)$, which is the regularity at which we aim to construct our solutions. Consequently, we do not use the randomness to overcome ill-posedness for low regularity data, but rather we use randomization to construct a rough but highly structured perturbation of the original log-log dynamics.

Remark 1.5. We also want to mention the recent interesting work [43], which implies that blowup solutions to the one-dimensional mass critical NLS cannot be seen by the Gibbs measure when incorporating a mass truncation less than or equal to the mass of the ground state. There are several differences[4] between our work and [43], but the most important conceptual difference is that the blowup solutions studied in the current article all have mass strictly above the mass of the ground state.

1.2. Comparison with previous results

1.2.1. Log-log blowup in H^1. We start with a quick review of the works of Merle and Raphaël [36–38, 40]. Let us focus on $H^1(\mathbb{R}^2)$ solutions u to (1.1) with negative energy, zero momentum,[5] and with mass slightly above that of the ground state Q; see (1.4). Via a variational argument and modulation theory, one can establish a geometric decomposition for the solution, given by

$$u(t,x) = \frac{1}{\lambda(t)}(\widetilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{i\gamma(t)}, \tag{1.12}$$

where $\widetilde{Q}_b$ is a certain elliptic object which is a modification of Q (see (2.18)) such that certain orthogonality conditions given in (3.21)–(3.24) below hold, and b and ϵ are a priori small.

One may then reduce the study of (1.1) to that of the evolution[6] of $\epsilon(t,x)$ and the parameters $b(t), \lambda(t), \gamma(t), x(t)$. It turns out that one should study this system in a rescaled time variable s rather than the original time variable, where

$$\frac{dt}{ds} = \lambda^2. \tag{1.13}$$

We note that λ dictates the blowup rate, and the parameter b dictates the evolution[7] of λ in the sense that $b \sim -\lambda_s/\lambda$.

[4] For example, their random data is centered, in some sense, at 0, whereas our random initial data is a random perturbation centered at the ground state up to some symmetry. Also, our solution is at L^2_x regularity, whereas their solutions are relatively regular, in $H^{1/2-}$.

[5] One can always perform a Galilean transformation to set the momentum to zero, which does not change the mass and does not increase the energy.

[6] Heuristically we now have five unknowns, $\epsilon, b(t), \lambda(t), x(t), \gamma(t)$, and five equations, (1.1), (3.21)–(3.24). Thus, one may expect the system is well determined.

[7] Or more precisely, $\widetilde{Q}_b$ is constructed in a such a way that $b \sim -\lambda_s/\lambda$.

A key estimate in the analysis of Merle and Raphaël is a local virial estimate[8],

$$b_s \geq H(\epsilon) - 2\lambda^2 E - \Gamma_b^{1-C\eta} + o(1)\left(\int |\nabla\epsilon|^2 + \epsilon^2 e^{-|y|}\right), \tag{1.14}$$

where Γ_b is a certain quantity which we define in (2.20) below, that satisfies

$$e^{-(1+C\eta)\frac{\pi}{b}} \leq \Gamma_b \leq e^{-\frac{\pi}{b}(1-C\eta)}. \tag{1.15}$$

for $C\eta \ll 1$. One hopes to deduce from (1.14) that

$$b_s \geq -\Gamma_b^{1-C\eta}. \tag{1.16}$$

Note that formula (1.16) is closely related to the sharp upper bound of the log-log blowup. The main point is H in (1.14) is some quadratic form, which will be coercive, dominating

$$\int |\nabla\epsilon|^2 + \epsilon^2 e^{-|y|}$$

up to six "bad" directions. Four bad directions will be handled via orthogonality of the modulation parameters,[9] the other two are handled by energy and momentum conservation. We also remark here that when E is negative, the term $-2\lambda^2 E$ in (1.14) will not pose any problems for the analysis. In other words, heuristically, one only needs control of the positive part of E.

Remark 1.6. The estimate (1.16) already implies the sharp upper bound on the blowup rate for the log-log dynamics, which coincides with the direct scaling lower bound up to a double logarithm. To derive the sharp lower bound, one needs to introduce a truncated object $\tilde{\zeta}_b$, defined in (2.22) to further sharpen the analysis, see Sections 2.4 and 6.

We note that (1.16) is enough to drive the dynamics into a regime where

$$\lambda \ll e^{-e^{\Gamma_b^{-c}}}. \tag{1.17}$$

In this regime, the crucial observation in [20] is that when λ is small compared to b, one does not need the negative energy assumption anymore since $\lambda^2|E|$ can be treated as a small perturbation. A similar mechanism can also be applied to momentum, i.e. one does not require the strict zero momentum condition.

1.3. H^s stability of log-log blowup

The study of the log-log dynamics in $H_x^s(\mathbb{R}^2)$ can be split into two stages. The first stage establishes rigidity of the dynamics in the sense that the solution will be driven towards some special, well-prepared initial data with an almost explicit form. The second stage

[8]Note that this estimate only involves local L^2 information.

[9]In practice, some extra is cancellation is needed, since one of the orthogonality conditions for the modulation parameters does not directly compensate for one of the bad directions associated with H.

establishes that for such well-prepared data, its evolution can be understood via a bootstrap argument; see [45]. Though both stages will rely on the same crucial ingredients from the analysis of Merle and Raphaël, the dynamics in the second stage is better understood since one can argue explicitly by bootstrap.

The Cauchy problem (1.1) is locally well-posed in $H_x^s(\mathbb{R}^2)$ for any $s > 0$. Thus, to prove $H_x^s(\mathbb{R}^2)$ stability of log-log blowup dynamics is equivalent to proving that for those well-prepared initial data whose evolution can be characterized by the bootstrap estimates, the evolution is stable under $H_x^s(\mathbb{R}^2)$ perturbations. This fact is established in the work of Colliander and Raphaël [20]. One crucial observation and heuristic is that since the solution u is of the form

$$u = \frac{1}{\lambda(t)} h\left(\frac{x}{\lambda(t)}\right), \quad \|h(t)\|_{H^1} \sim 1,$$

in the $H_x^1(\mathbb{R}^2)$ case, one may expect that in the $H_x^s(\mathbb{R}^2)$ case the solution has a similar structure, with "quantitative energy bounds"[10]

$$E(u) \sim \frac{1}{\lambda^{2-2s}}.$$

Recall that in (1.14), the term $E(u)$ has been multiplied by λ^2 in the analysis, and will be formally of size λ^{2s}. When λ satisfies (1.17), this term can be treated perturbatively provided $s > 0$. This also explains why $s = 0$ is conceptually different from the case $s > 0$.

Unsurprisingly, a main challenge in the analysis of [20] is that since u is not in $H^1(\mathbb{R}^2)$ anymore, the energy $E(u)$ is not well-defined (indeed, otherwise, it would be bounded by a constant). To overcome this difficulty, one employs the I-method, introduced by Colliander, Keel, Staffilani, Takaoka and Tao [18], a ubiquitous method in the study of dispersive PDEs which exploits energy conservation for low regularity data, and which is philosophically similar (although practically not completely equivalent) to the high-low method of Bourgain [6]. We note that it may be surprising that one can apply the I-method for all $s > 0$, whereas typically, such computations only work for certain $s > s_0$. Broadly speaking, this is because one has a good *a priori* understanding of the log-log asymptotics.

To briefly sketch the strategy of [20], one still considers the ansatz

$$u(t,x) = \frac{1}{\lambda(t)} (\tilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right) e^{-i\gamma(t)}. \tag{1.18}$$

One applies the time-dependent operator $I_{N(t)}$ which truncates the high frequency part of the solution above $N(t) = \lambda(t)^{-(1+)}$. One then aims to study the evolution of $I_{N(t)}u(t)$, and to prove that the *positive part* of energy $E(I_{N(t)}u(t))$ is controlled by λ^{-2+2s} (formally speaking) via the I-method. In particular, one must establish that this energy cannot be too large and positive, although it can be very negative.[11] A key observation in [20] is that the I-method is compatible with log-log bootstrap scheme; see also [45].

[10] We record this to give some intuition, but in practice, one needs to perform a frequency truncation to discuss the energy of u.

[11] It is also emphasized in [20] that the negative part of the energy always drives the solution to blowup.

1.4. Random L^2 perturbations

The approach in [20] breaks down for general L_x^2 perturbations. In this article, we will use randomized $L_x^2(\mathbb{R}^2)$ data f^ω, defined in (1.9), to perturb essentially the same well-prepared data as in [20]. For the solutions $u(t,x)$ to (1.1), we use the ansatz

$$u(t,x) = a(t,x) + F(t,x), \tag{1.19}$$

where

$$a = \frac{1}{\lambda(t)}(\widetilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{-i\gamma(t)} \quad \text{and} \quad F = e^{it\Delta} f^\omega.$$

Our a will behave as the full solution u of [20].

The study of dispersive PDEs via a probabilistic approach was initiated by Bourgain [4, 5] for the periodic nonlinear Schrödinger equation in one and two space dimensions, building upon the constructions of invariant measures by Glimm and Jaffe [26] and Lebowitz, Rose and Speer [32]. Such questions were further explored by Burq and Tzvetkov [13, 14] in the context of the cubic nonlinear wave equation on a three-dimensional compact Riemannian manifold. There has since been a vast body of research where probabilistic tools are used to study many nonlinear dispersive or hyperbolic equations at supercritical regularities: see for instance the works [1, 8, 11, 19, 21, 41] as well as references and discussion therein.

Certain global-in-time random data results in the compact setting which rely on invariant measures work equivalently in the focusing and defocusing cases [4].[12] However, in the absence of an invariant measure, the vast majority of existing large data probabilistic results treat only the *defocusing* nonlinear Schrödinger and wave equations: see for instance [9, 15, 23, 30, 33, 34, 42, 46] and references therein. We note that analogously to the deterministic theory, one may occasionally obtain "small data" type probabilistic results in the focusing setting (see for instance [33]), although these are consequences of the local theory and do not relate to the large data probabilistic techniques.

There are two recent works in particular which treat the focusing problem with random initial data, outside the small data or local-in-time regimes. The first is work of Kenig and Mendelson [29] which studies the probabilistic stability of the soliton for the energy critical nonlinear wave equation on $\mathbb{R}^3$. In that work, the authors produce with high probability a family of radial perturbations of the soliton which give rise to global forward-in-time solutions of the focusing nonlinear wave equation that scatter after subtracting a dynamically modulated soliton. The proof relies on a new randomization procedure using distorted Fourier projections associated to the linearized operator around a fixed soliton. Another work, also in the context of nonlinear wave equations, is recent work of Bringmann [10] on the probabilistic stability of the ODE blowup for the quintic nonlinear wave equation on $\mathbb{R}^3$. The proof in the latter paper relies on probabilistic

[12] Even so, this only holds in dimension $d = 1$, since in dimension $d \geq 2$, the construction of the measure fails for focusing nonlinearities: see e.g. [12].

Strichartz estimates in similarity coordinates, and in particular does not require a randomization adapted to the blowup solution. We note that in the current work, our ansatz (1.19) separates the free evolution as opposed to a linearized evolution as in [29]. The reason for this is that the current work handles the finite time blowup problem whereas [29] concerned the infinite time dynamics centered at a modulated, but nonconcentrating, soliton. It is the analysis of the asymptotic dynamics which requires one to adapt the randomization to the linearized operator.

As in [29] and [10], in the current work we are in the large data yet perturbative regime. We note, however, that the geometric blowup we treat is quite distinct from the ODE blowup handled in [10]. We leverage random data techniques to establish a bootstrap result, stated in Lemma 3.6 below. In the present work, compared to previous random data works, we do not need to use probabilistic improvements to overcome issues with deterministic well-posedness. Indeed, as mentioned previously, (1.1) is deterministically locally well-posed in $L_x^2(\mathbb{R}^2)$ via Strichartz estimates. However, we leverage the random data in two novel ways: first, we use it to obtain precise quantitative control on the well-posedness estimates for the bootstrap scheme, and second, we use the randomness to achieve the endpoint estimates[13] for the I-method computation mentioned above. The former estimates are achieved in the spirit of the work of Bourgain [7], adapted to $\mathbb{R}^2$. One difference between our work and Bourgain's is that since our initial data lies at $L_x^2(\mathbb{R}^2)$ regularity, there is no need to "Wick-order" the nonlinearity[14].

In the I-method computation, we exploit improved probabilistic estimates for the free evolution of the random initial data in a novel manner; mainly we use the fact that they can be made uniformly small in time to close the I-method estimates at the endpoint. Heuristically, one can view the free evolution of the random data as being not only equidistributed in space for a fixed time, but also roughly equidistributed in time. Hence, from the point of view of the time scales associated to log-log blowup, the free evolution of the random data can be thought of as a source term which makes increasingly small contributions to the dynamics, and thus preserves the blowup mechanism. Finally, we remark that although a priori a and F are both $\mathcal{O}(1)$ in $L_x^2(\mathbb{R}^2)$, since the randomized initial data is nonconcentrated, the L^2-pairing of the singular part, a, with the free evolution of the random data, F, will give a power of λ, and hence all such terms will be of perturbative nature in the modulation argument since λ is so small. We will carry out a more thorough discussion of the proof in Section 1.5.

[13]More precisely, though the I-method computation evolves terms with endpoint regularity, we apply the probalistic techiniques to prove they will behave like non-endpoint elements, thus the problem is still subcritical rather than critical. It should be noted that both I-method and random data analysis are of subcritical nature, though in those problems, the original critical regularity may not be critical in the usual sense.

[14]The precise form of such an ordering is not as obvious in the Euclidean setting, but a continuous version of Wick ordering is indeed possible.

1.5. An overview for the proof of the bootstrap lemma

We conclude the introduction by presenting an overview for the proof of the main bootstrap result, stated in Lemma 3.6. Two ingredients are required to close the bootstrap:

- Under the bootstrap assumption, the dynamics can be viewed as a perturbation of the log-log blowup dynamics.
- Log-log blowup dynamics can upgrade the bootstrap assumptions to bootstrap estimates.

We will focus on the first ingredient, since the second part essentially follows from earlier works, in particular [20], building on the earlier works [40, 45]. There are three main factors which ensure the dynamics can be viewed as a perturbation of the log-log dynamics:

- According to the bootstrap assumptions, we have $\lambda \ll b$, and essentially all terms of the form λ^σ for $\sigma > 0$ may be treated as a perturbation. In particular, if one pairs the linear evolution of the randomized data F with terms of the form $\frac{1}{\lambda}h(x/\lambda)$ such that h is somehow localized, one obtains a perturbative term.
- The bootstrap assumption $t_{k+1} - t_k \lesssim k\lambda(t_k)^2$ gives good control on how many local well-posedness (LWP) intervals we have throughout the analysis, and in particular, in every LWP interval $[a, b]$ such that $\lambda(t) \sim 2^{-k}$ and $|b - a| \sim \lambda^2(t_k)$, we can establish probabilistic well-posedness, based on the bootstrap assumption (3.32).
- Finally, based on the probabilistic well-posedness in every LWP interval, one can perform an I-method type energy estimate combined with random data estimates to obtain good control on $E(I_{N(t)}u)$ (or more precisely, the positive part of this quantity), which will ensure the log-log dynamics persists. One also needs to control $P(I_N u)$, but this is relatively easier.

The key ingredients in the current article are the development of suitable probabilistic well-posedness in every LWP interval, and the derivation of good energy estimates for $E(I_N u)$. One may compare these ingredients to those in [20], in which the usual $H^s_x(\mathbb{R}^2)$ well-posedness is used in every LWP interval, and a more classical version of the I-method is applied. As noted earlier, although there are certainly crucial differences between the current work and [20], we fundamentally rely on the observation from that work that the I-method is compatible with the log-log bootstrap regime.[15]

1.6. Organization of paper

In Section 2, we introduce some probabilistic and deterministic preliminaries. In Section 3 we describe the initial data and introduce the bootstrap assumptions. We will elaborate on the probabilistic estimates and I-method type computation in Section 4. In Section 5, we

[15] In some sense, all I-method arguments rely on good control of the number of LWP intervals, and a good understanding of well-posedness estimates within every LWP interval.

will provide a relatively detailed sketch of how such an energy estimate plus the log-log dynamics close the bootstrap scheme in Section 6.

1.7. Notation

We use $\Lambda := 1 + y \cdot \nabla$ to denote the generator of the $L^2_x(\mathbb{R}^2)$ scaling. When we write $f = f_1 + if_2$, we implicitly mean $f_1 := \Re f$ and $f_2 := \Im f$. We denote by $C > 0$ an absolute constant which only depends on fixed parameters and whose value may change from line to line. We write $X \lesssim Y$ to indicate that $X \leq CY$, and $X \sim Y$ if $X \lesssim Y \lesssim X$. Moreover, we write $X \lesssim_\nu Y$ to indicate that the implicit constant depends on a parameter ν and we write $X \ll Y$ if the implicit constant should be regarded as small. We will write $c+$ to denote $c + \varepsilon$ for an arbitrary $\varepsilon > 0$, and similarly for $c-$. We also use the notation $\langle x \rangle := (1 + x^2)^{1/2}$.

2. Preliminaries

2.1. I-operator

Following [18, 20], let $0 < s < 1$ and let $m : \mathbb{R}^+ \to \mathbb{R}^+$ be a smooth, monotone function which satisfies $m(|\xi|) = 1$ for $0 \leq |\xi| \leq 1$, and $m(|\xi|) = |\xi|^{s-1}$ for $|\xi| \geq 2$. Let $N \gg 1$ and define

$$m_N(\xi) = m(|\xi|/N),$$

and note that

$$m_N(|\xi|) = \begin{cases} 1, & |\xi| < N, \\ (N/|\xi|)^{1-s}, & |\xi| > 2N. \end{cases} \tag{2.1}$$

The operator I_N is the Fourier multiplier associated to m_N:

$$\widehat{I_N f}(\xi) = m_N(\xi)\hat{f}(\xi)$$

and we note that

$$\|f\|_{H^s} \lesssim \|I_N \langle D \rangle f\|_{L^2} \lesssim N^{1-s} \|f\|_{H^s}.$$

Remark 2.1. The operator I_N is also strong-type (p, p) for all $1 \leq p \leq \infty$, uniformly in N.

2.2. Strichartz estimates

We recall the classical Strichartz estimates, which play an important role in the local theory of NLS.

Definition 2.2 (Admissible pairs). For $d \geq 1$ we say a pair of exponents (q, r) is *Schrödinger admissible* if

$$\frac{2}{q} + \frac{d}{r} = \frac{d}{2}, \quad 2 \leq q, r \leq \infty, \quad \text{and} \quad (d, q, r) \neq (2, 2, \infty). \tag{2.2}$$

For a fixed spacetime slab $I \times \mathbb{R}^d$, we define the Strichartz norm

$$\|u\|_{S(I)} := \sup_{(q,r)\text{ admissible}} \|u\|_{L_t^q L_x^r(I\times\mathbb{R}^d)}. \tag{2.3}$$

We let $S(I)$ denote the closure of all test functions under this norm, and let $N(I)$ denote its dual.

Remark 2.3. In dimension $d = 2$, the supremum must actually be restricted to a closed subset to avoid the inadmissible endpoint.

Proposition 2.4 (Strichartz estimates [16,28,48]). *Let $0 \le s \le 1$, let I be a compact time interval, and let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be a solution to the forced Schrödinger equation*

$$iu_t + \Delta u = F.$$

Then for any $t_0 \in I$, we have

$$\big\||\nabla|^s u\big\|_{S(I)} \lesssim \|u(t_0)\|_{\dot{H}_x^s} + \big\||\nabla|^s F\big\|_{N(I)}.$$

Proposition 2.5 (Bilinear Strichartz estimates, cf. [6]). *Let f_1, f_2 be $L_x^2(\mathbb{R}^2)$ functions, and let $N \ge M$. Then*

$$\|e^{it\Delta}P_N f_1 e^{it\Delta}P_M f_2\|_{L^2_{t,x}} \lesssim (M/N)^{1/2}\|f_1\|_{L_x^2}\|f_2\|_{L_x^2}. \tag{2.4}$$

We now turn to the definition of $X^{s,b}$ spaces, [3]:

Definition 2.6. The space $X^{s,b}(\mathbb{R}\times\mathbb{R}^d)$ is the closure of test functions under the norm

$$\|u\|_{X^{s,b}(\mathbb{R}\times\mathbb{R}^d)} := \|\langle\xi\rangle^s\langle\tau - |\xi|^2\rangle\widehat{u}(\xi,\tau)\|_{L^2_{\xi,\tau}}.$$

Recall that $X^{s,b}$ embeds into $C_t^0 H_x^s$ for $b > 1/2$. The restricted version of the space on $[-\delta,\delta]\times\mathbb{R}^d$ is defined by

$$\|u\|_{X^{s,b,\delta}} := \inf\{\|\widetilde{u}\|_{X^{s,b}(\mathbb{R}\times\mathbb{R}^d)} \,:\, \widetilde{u}\big|_{[-\delta,\delta]} = u\}.$$

We recall that free solutions lie in $X^{s,b}$ *locally* in time but not globally. An important property of $X^{s,b}$ spaces is the following:

Lemma 2.7. *Let Y be a Banach space of functions on $\mathbb{R}\times\mathbb{R}^d$ with the property that*

$$\|e^{it\tau_0}e^{it\Delta}f\|_Y \lesssim \|f\|_{H_x^s}$$

for all $f \in H^s$ and $\tau_0 \in \mathbb{R}$. Then

$$\|u\|_Y \lesssim_b \|u\|_{X^{s,b}(\mathbb{R}\times\mathbb{R}^d)}.$$

We will also need a multilinear transfer principle for $X^{s,b}$ spaces:

Proposition 2.8 (Transfer principle, cf. [27]). *Let $b > 1/2$, $Y = L_t^q L_x^r$ for $1 \le p, q \le \infty$ and $\mathcal{T}$ a k-linear operator such that*

$$\|\mathcal{T}(e^{it\Delta} f_1, \dots, e^{it\Delta} f_k)\|_Y \lesssim \prod_{j=1}^{k} \|f_j\|_{H_j^s}.$$

Then

$$\|\mathcal{T}(u_1, \dots, u_k)\|_Y \lesssim \prod_{j=1}^{k} \|u_j\|_{X^{s_j,b}}.$$

We will use the transfer principle repeatedly throughout our estimates in order to combine Strichartz estimates with $X^{s,b}$ spaces.

2.3. Random data preliminaries

Here we collect some of the random data results which we will use. Recall $F = e^{it\Delta} f^\omega$, where f^ω has been defined in (1.9). We begin with the following ℓ^∞ Gaussian bound:

Lemma 2.9. *For every $\varepsilon > 0$, there exist $C, c > 0$ such that*

$$\mathbb{P}\Big(\bigcup_{n \in \mathbb{Z}^2} \{\langle n\rangle^{-\varepsilon} |g_n(\omega)| > K\}\Big) \le C e^{-cK^2}. \tag{2.5}$$

Next we record a standard probabilistic estimate.

Lemma 2.10 ([13, Lemma 3.1]). *Let $\{g_n\}_{n=1}^\infty$ be a sequence of complex-valued independent identically distributed (iid) mean-zero Gaussian random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then there exists $C > 0$ such that for every $p \ge 2$ and every $\{c_n\}_{n=1}^\infty \in \ell^2(\mathbb{N}; \mathbb{C})$, we have*

$$\Big\|\sum_{n=1}^{\infty} c_n g_n(\omega)\Big\|_{L_\omega^\rho} \le C\sqrt{\rho}\Big(\sum_{n=1}^{\infty} |c_n|^2\Big)^{1/2}.$$

We will also use the following variant of [50, Lemma 4.5] to bound the probability of certain subsets of the probability space.

Lemma 2.11. *Let F be a real-valued measurable function on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Suppose that there exist $\alpha > 0$, $N > 0$, $k \in \mathbb{N} \setminus \{0\}$ and $C > 0$ such that for every $\rho \ge \rho_0$,*

$$\|F\|_{L_\omega^\rho} \le C N^{-\alpha} \rho^{k/2}. \tag{2.6}$$

Then, there exist C_1, and δ depending on C and ρ_0, such that for $K > 0$,

$$\mathbb{P}(\omega \in \Omega : |F(\omega)| > K) \le C_1 e^{-\delta N^{2\alpha/k} K^{2/k}}. \tag{2.7}$$

Lemma 2.12. *Let* $1 \le p, q < \infty$. *Then for all* $\rho \ge \max(p,q)$ *we have*

$$\|F\|_{L^\rho_\omega L^q_t L^p_x([0,1]\times\mathbb{R}^2)} \lesssim_\rho \|f\|_{L^2_x}. \tag{2.8}$$

In particular, there exist $C, c > 0$ *such that*

$$\mathbb{P}(\|F\|_{L^q_t L^p_x([0,1]\times\mathbb{R}^2)} > K) \le Ce^{-cK^2/\|f\|^2_{L^2_x}}. \tag{2.9}$$

Proof. We use Minkowski's inequality and Lemma 2.10 to estimate

$$\begin{aligned} \|F\|_{L^\rho_\omega L^q_t L^p_x([0,1]\times\mathbb{R}^2)} &\le \|F\|_{L^q_t L^p_x([0,1]\times\mathbb{R}^2) L^\rho_\omega} \\ &\lesssim \Big(\sum_k \|e^{it\Delta} P_k f\|^2_{L^q_t L^p_x([0,1]\times\mathbb{R}^2)}\Big)^{1/2}. \end{aligned} \tag{2.10}$$

The result then follows from Hölder's inequality and the unit-scale Bernstein inequality, while the estimate on the probability follows from Lemma 2.11. ∎

Remark 2.13. Essentially repeating the proof of the previous lemma, one may apply the Minkowski inequality and the ℓ^2 summability (1.11) to improve (2.8) to

$$\Big\|\Big(\sum_N \|P_N F\|^2_{L^p_{t,x}([0,1]\times\mathbb{R}^2)}\Big)^{1/2}\Big\|_{L^\rho_\omega} \lesssim_\rho 1. \tag{2.11}$$

where $N \ge 1$ ranges over all dyadic integers. Via interpolation with the $L^\infty_t L^2_x$ bound and Hölder's inequality in time, up to an exceptional set of exponentially small probability, one has, for $2 \le q \le p < \infty$,

$$\|F\|_{L^p_t L^q_x([0,1]\times\mathbb{R}^2)} \lesssim 1, \tag{2.12}$$

$$\sum_N \|P_N F\|^2_{L^p_t L^q_x([0,1]\times\mathbb{R}^2)} \lesssim 1. \tag{2.13}$$

Finally, we will need a multilinear Gaussian estimate. We state a slightly simplified version of this estimate compared to the reference since it will suffice for our purposes.

Lemma 2.14 (Cf. [49, Proposition 2.4]). *Let* $\{g_n\}$ *be iid mean-zero Gaussian random variables, and let*

$$* = \{n_1, n_2, n_3 \in \mathbb{Z}^2 \,:\, n_2 \ne n_1, n_3\}.$$

Consider

$$G(\omega) = \sum_* c(n_1,n_2,n_3) g_{n_1}(\omega)\overline{g}_{n_2}(\omega) g_{n_3}(\omega) \tag{2.14}$$

where $c(n_1,n_2,n_3)$ *are complex numbers. Then there exist* $C, c > 0$ *such that*

$$\mathbb{P}\{|G| > K\|G\|_{L^2_\omega}\} \le Ce^{-cK^2}. \tag{2.15}$$

2.4. Elliptic objects

We recall, as detailed in the introduction, that analysis of Merle and Raphaël [38] for log-log blowup solutions begins with a geometric decomposition of the solution u, given by

$$u(t,x) = \frac{1}{\lambda(t)}(\widetilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{i\gamma(t)}, \tag{2.16}$$

where $\widetilde{Q}_b$ is a certain elliptic object which is a modification of Q and where ϵ is a priori small in $H^1_x(\mathbb{R}^2)$. Schematically, this implies that a scaled and translated version of the ground state Q is a good approximation for u, up to modulation. However, in practice, one considers a modification of Q to capture the sharp log-log blowup dynamics [36, 37, 40]. This modification of Q relies on certain elliptic objects, Q_b, $\widetilde{Q}_b$, ζ_b and $\widetilde{\zeta}_b$, which we describe in this subsection. We will not list all the properties of these objects, which are indeed crucial to log-log analysis but will not explicitly be used in this article, since we rely on previous results which establish the existence of such solutions in the energy space. Instead, we focus only on the properties which are most relevant to the current work and refer to [36, 37, 40] for more details. One may refer to [40, Proposition 1 and Lemma 2] for further details.

Throughout this subsection, b and η will be used to denote small positive numbers, C will denote a universal constant, and one should have in mind that $C\eta \ll 1$. We let

$$R_b := \frac{2}{|b|}\sqrt{1-\eta}.$$

Let Q_b be a modification of Q which solves

$$\begin{cases} \Delta Q_b - Q_b + ib\Lambda Q_b + |Q_b|^2 Q_b = 0, \\ Q_b e^{ib|y|^2/4} > 0 \quad \text{in } B_{R_b}, \\ Q_b(R_b) = 0. \end{cases} \tag{2.17}$$

Now, let $R_b^- := \sqrt{1-\eta}\, R_b \sim b^{-1}$ be a constant slightly smaller than R_b, and let $\phi_b(x)$ be a smooth cut-off function with $\phi_b(x) \equiv 1$ on $|x| \le R_b^-$ and $\phi_b(x) \equiv 0$ for $|x| \ge R_b$. We define $\widetilde{Q}_b$ to be the cut-off version of Q_b, namely $\widetilde{Q}_b = \phi_b Q_b$, and we let

$$\Delta\widetilde{Q}_b - \widetilde{Q}_b + ib\Lambda\widetilde{Q}_b + |\widetilde{Q}_b|^2\widetilde{Q}_b =: -\Psi_b. \tag{2.18}$$

Note that $\widetilde{Q}_b(x)$ decays exponentially as $|x| \to \infty$, thus asymptotically the nonlinearity of (2.18) vanishes.

Following the work of Merle and Raphaël [40], one introduces the tail ζ_b, which is the unique radial solution to

$$\begin{cases} \Delta\zeta_b - \zeta_b + ib\Lambda\zeta_b = \Psi_b, \\ \zeta_b \in \dot{H}^1_x(\mathbb{R}^2). \end{cases} \tag{2.19}$$

It turns out that ζ_b just misses $L^2_x(\mathbb{R}^2)$, or more precisely, if we define

$$\Gamma_b := \lim_{|y|\to\infty} y^2|\zeta_b|^2, \tag{2.20}$$

then this limit exists and we have

$$e^{-(1+C\eta)\frac{\pi}{b}} \le \Gamma_b \le e^{-\frac{\pi}{b}(1-C\eta)}. \tag{2.21}$$

The quantity Γ_b appears frequently in the log-log blowup analysis, and this scale plays a crucial role. A useful heuristic to keep in mind is that all terms of size Γ_b^{1+} are acceptable. For example, if one modifies ζ_b into ζ_b' so that $\|\widetilde{\zeta}_b - \widetilde{\zeta}_b'\|_{H^1} \lesssim \Gamma_b^{1+}$, then, heuristically, there is no difference between those two terms in the log-log analysis.

To overcome the failure of $L^2_x(\mathbb{R}^2)$ integrability of ζ_b, we introduce a cut-off version of this object, denoted by ζ_b, as follows. Let ψ be a bump function localized at $|x| \lesssim 1$ and let a be a small number. Let

$$A = A_b := e^{a\pi/b}, \quad \psi_A(x) := \psi(x/A),$$

and let

$$\widetilde{\zeta}_b = \psi_A \zeta_b. \tag{2.22}$$

Note that $\Gamma_b^{-a/2} \le A_b \le \Gamma_b^{-3a/2}$ and

$$\int |\widetilde{\zeta}_b|^2 \le \Gamma_b^{1-C\eta}. \tag{2.23}$$

One also records

$$\Delta\widetilde{\zeta}_b - \widetilde{\zeta}_b + ib\Lambda\widetilde{\zeta}_b =: \Psi_b + F_b. \tag{2.24}$$

The crucial fact about the tails, used essentially in [40, (4.20)], is that

$$-\Re(\widetilde{\zeta}_b, \Lambda F_b) \ge c\Gamma_b. \tag{2.25}$$

We conclude this section by listing some useful estimates for $\widetilde{Q}_b$. Most of the time, however, it will be enough to think of it as a function which decays exponentially, uniformly in b.

(1) One has

$$|E(\widetilde{Q}_b)| \lesssim \Gamma_b^{1-C\eta}, \quad P(\widetilde{Q}_b) = 0. \tag{2.26}$$

(2) $\widetilde{Q}_b$ is uniformly close to Q, and

$$\|e^{(1-\eta)\theta(|b|\,|y|)/|b|}(\widetilde{Q}_b - Q)\|_{C^3} \xrightarrow{b\to 0} 0, \tag{2.27}$$

where

$$\theta(r) = \mathbf{1}_{\{0\le r\le 2\}} \int_0^r \sqrt{1-z^2/4}\,dz + \mathbf{1}_{\{r>2\}} \frac{\theta(2)}{2} r, \tag{2.28}$$

and $\theta(2) = \pi/2$.

(3) One has the following non-degeneracy with parameter b:

$$\left\| e^{(1-\eta)\theta(|b|\,|y|)/|b|} \left(\frac{\partial}{\partial_b} \widetilde{Q}_b + i \frac{|y|^2}{4} Q \right) \right\| \xrightarrow{b\to 0} 0. \tag{2.29}$$

(4) $\widetilde{Q}_b$ has strictly supercritical mass and

$$\|\widetilde{Q}_b\|^2 - \|Q\|_2^2 \sim b^2. \tag{2.30}$$

3. Preparation of initial data and setting up the bootstrap

In this section, we describe the necessary steps in order to set up the main bootstrap lemma and prove the main theorem.

3.1. Description of initial data and statement of main results

Recall that we consider a randomized $L^2_x(\mathbb{R}^2)$ function f^ω, given by

$$f^\omega(x) = \sum_k \int f_k g_k(\omega) \psi_k(\xi) e^{ix\xi}\, d\xi \tag{3.1}$$

where $\{g_k\}_{k\in\mathbb{Z}^2}$ are iid mean-zero complex Gaussian random variables, and where ψ_k and f_k are defined in (1.6) and (1.7). We recall that we assume that the f_k satisfy the decay condition

$$|f_k| \le 1/|k|, \quad k \neq 0, \tag{3.2}$$

and normalization

$$\sum_k |f_k|^2 = 1. \tag{3.3}$$

We will use F to denote the linear evolution of the random data f^ω, that is,

$$F(t,x) = F^\omega(t,x) = e^{it\Delta} f^\omega. \tag{3.4}$$

We let a_0 be the well-prepared initial data, given by

$$a_0 = \frac{1}{\lambda_0} (\widetilde{Q}_{b_0} + \epsilon_0) \left(\frac{x - x_0}{\lambda_0} \right), \tag{3.5}$$

and one may, without loss of generality, take $x_0 = 0$. We make the following assumptions which will ensure that we fall in the bootstrap regime of the log-log dynamics:

- smallness of b_0:

$$0 < b_0 \ll 1, \tag{3.6}$$

- smallness of λ_0:

$$0 < \lambda_0 \le e^{-e^{4\pi/(5b_0)}}, \tag{3.7}$$

- smallness of extra mass:

$$\|\epsilon_0\|_{L^2_x} \ll 1, \tag{3.8}$$

- H^1 smallness of ϵ_0:

$$\int |\nabla \epsilon_0|^2 + |\epsilon_0|^2 e^{-|y|} \le \Gamma_{b_0}^{4/5}, \tag{3.9}$$

- control of energy and momentum:

$$\lambda_0^{1/2}|E(a_0)| \le 1, \tag{3.10}$$

$$\lambda_0^{1/2}|P(a_0)| \le 1, \tag{3.11}$$

and the following four orthogonality conditions:

$$(\epsilon_{1,0}, |y|^2\Sigma_{b_0}) + (\epsilon_{2,0}, |y|^2\Theta_{b_0}) = 0, \tag{3.12}$$

$$(\epsilon_{1,0}, y\Sigma_{b_0}) + (\epsilon_{2,0}, y\Theta_{b_0}) = 0, \tag{3.13}$$

$$-(\epsilon_{1,0}, \Lambda\Theta_{b_0}) + (\epsilon_{2,0}, \Lambda\Sigma_{b_0}) = 0, \tag{3.14}$$

$$-(\epsilon_{1,0}, \Lambda^2\Theta_{b_0}) + (\epsilon_{2,0}, \Lambda^2\Sigma_{b_0}) = 0, \tag{3.15}$$

where

$$\epsilon_0 = \epsilon_{1,0} + i\epsilon_{2,0}, \quad \tilde{Q}_{b_0} = \Sigma_{b_0} + i\Theta_{b_0} \tag{3.16}$$

and (f, g) denotes the real L^2_x inner product. We remark that such initial data a_0 are easy to construct by the work of Merle and Raphaël [40]. Indeed, one simply finds $H^1_x(\mathbb{R}^2)$ initial data, with non-positive energy and mass slightly above that of the ground state, and evolves it under the flow of (1.1) until it is close enough to the blowup time.

Here and below, we assume that f and a_0 satisfy the above conditions. We are now prepared to state our main result.

Theorem 3.1. *Fix f satisfying the above conditions. There exists a universal constant $\lambda_0^* > 0$ such that for all $0 < \lambda_0 < \lambda_0^*$, there exists a subset $\Sigma \subset \Omega$ and constants $C, c > 0$ such that*

$$\mathbb{P}(\Sigma) \ge 1 - Ce^{-1/\lambda_0^c},$$

and for all $\omega \in \Sigma$, there exists a solution $u(t,x)$ to (1.1) with initial data $u_0 = a_0 + f^\omega$ which will blow up in finite time $0 < T = T_\omega \ll 1$ according to the log-log law in the following sense: there are two small, fixed positive numbers s, δ such that

$$u(t,x) = a(t,x) + F(t,x), \quad a(t,x) = \frac{1}{\lambda(t)}(\tilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right), \quad F = e^{it\Delta} f^\omega \tag{3.17}$$

and

$$\lambda(t)^{-1} \sim \frac{\sqrt{\log|\log|T-t||}}{\sqrt{T-t}}, \quad \|\tilde{Q}_b + \epsilon\|_{H^s} \sim 1, \tag{3.18}$$

and for some $N(t) = \lambda(t)^{-1-\delta}$,

$$\int |\nabla I_{N(t)\lambda(t)}\epsilon|^2 + |\epsilon|^2 e^{-|y|} \xrightarrow{t\to T} 0. \tag{3.19}$$

One may refer to Section 2.1 for the definition and properties of the I-operator.

Before proceeding, we make a remark about the "large probability" in the statement of Theorem 3.1.

Remark 3.2. The large probability in the statement of the theorem can be understood in two ways. If one fixes a_0, and studies the evolution of $a_0 + \alpha f^\omega$, then with probability $\geq 1 - e^{-1/\alpha^c}$ the conclusion of Theorem 3.1 holds provided α is sufficiently small. Alternatively, one may fix f, but consider λ_0 and b_0 sufficiently small, since the definition of a_0 is essentially given by λ_0 and b_0. Then the conclusion of Theorem 3.1 holds with probability $\geq 1 - e^{-1/\lambda_0^c}$. Note that the exceptional set that one needs to drop so that the conclusion of Theorem 3.1 holds, essentially depends on λ_0, given that λ_0 is already chosen sufficiently small[16] depending on b_0.

Theorem 3.1, as stated, can be interpreted as providing a construction of L_x^2 log-log blowup solutions. In [20], the extension from this construction with its additional quantitative information to general stability of log-log blowup solutions is achieved via the H_x^s local theory. In our case, however, we must rely on our improved probabilistic well-posedness instead of the classical L_x^2 local theory. We elaborate on this in the following remark.

Remark 3.3. Theorem 3.1, or more precisely, the proof of Theorem 3.1 implies stability of log-log blowup under certain random L_x^2 perturbations. More precisely, consider initial data $v_0 \in H^1$ with mass slightly above the ground state mass which leads to a log-log blowup. Note that v_0 may be far away from the form a_0 given by (3.5). We claim that if one perturbs v_0 with αf^ω, for α sufficiently small, the corresponding solution will still blow up (with high probability) according to the log-log law as in Theorem 3.1.

To see this, first note that if v is the solution to (1.1) with initial data v_0, then at time T_1, the solution $v(T_1)$ will enjoy the same properties and general form as a_0 (see (3.5)). Note that no matter how v evolves after time T_1, v is still well-behaved within $[0, T_1]$, specifically v has finite Strichartz norm within that time interval and one can thus apply perturbation theory purely via the local theory. If one considers the evolution of $v_0 + \alpha f^\omega$, if one simply applies the classical L_x^2 local theory, it is not enough to conclude. However, if one now applies our improved probabilistic well-posedness, established in Section 4, and lets $\tilde{v}$ be the solution to (1.1) with initial data $v_0 + \alpha f^\omega$, one will find that

$$\tilde{v}(T_1) = v(T_1) + \alpha e^{iT_1\Delta} f^\omega + h_\omega,$$

where, off an exceptional set of small probablity, h_ω is small[17] in H^s. One can then carry out the proof of Theorem 3.1 line by line to establish that such data will blow up in the manner described in Theorem 3.1, in particular according to the log-log law.

---

[16] In some sense, there are only two effective parameters in a, which are λ and b, and all the constraints for ϵ depend on b. Such a b must be chosen small enough. Then, λ needs to be chosen small enough according to b, i.e. $\lambda \leq \lambda^*(b)$.

[17] Depending on the smallness of α.

We also have the following two remarks.

Remark 3.4. At first glance, the statement of the main theorem may seem surprising, since one could choose a_0 so concentrated that one does not even need the $L^2_x(\mathbb{R}^2)$ smallness of f^ω, or alternatively smallness of α. One should still view the free evolution of the random data as a (small) perturbation (around a complicated object) since the requirement that a_0 be concentrated, together with the fact that randomized functions are equidistributed in space, still decouples these terms from one another, and thus the resulting interaction is still expected to be small.

Remark 3.5. Just as the $H^1_x(\mathbb{R}^2)$ case and the $H^s_x(\mathbb{R}^2)$ case for $s > 0$, one can study the convergence of the concentration point $x(t)$ to establish that the blowup point is well-defined, and one can prove (non)concentration properties of the radiation ϵ at the blowup point. We refer the interested readers to [20] since these arguments apply in an identical manner in our setting.

The dynamics described in the main theorem will be characterized by the bootstrap lemma in the next subsection.

3.2. *Bootstrap setup*

Let u be the solution to (1.1) with initial data u_0. We will use the ansatz

$$\begin{aligned} u(t,x) &= a(t,x) + F(t,x), \\ a(t,x) &= \frac{1}{\lambda(t)}(\widetilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{-i\gamma(t)}. \end{aligned} \tag{3.20}$$

Via continuity of the flow in L^2_x and our initial orthogonality conditions (3.12)–(3.15), we can ensure that, at least locally in $t \in [0, T_0]$, T_0 small, one has the orthogonality conditions

$$(\epsilon_1, |y|^2\Sigma_b) + (\epsilon_2, |y|^2\Theta_b) = 0, \tag{3.21}$$

$$(\epsilon_1, y\Sigma_b) + (\epsilon_2, y\Theta_b) = 0, \tag{3.22}$$

$$-(\epsilon_1, \Lambda\Theta_b) + (\epsilon_2, \Lambda\Sigma_b) = 0, \tag{3.23}$$

$$-(\epsilon_1, \Lambda^2\Theta_b) + (\epsilon_2, \Lambda^2\Sigma_b) = 0. \tag{3.24}$$

and furthermore

$$x(0) = x_0, \quad \lambda(0) = \lambda_0, \quad \gamma(0) = 0. \tag{3.25}$$

where we set $\epsilon = \epsilon_1 + i\epsilon_2$ and $\widetilde{Q}_b = \Sigma_b + i\Theta_b$.

We will focus on the evolution of a, which satisfies

$$i\partial_t a + \Delta a = -|a + F|^2(a + F) = -|a|^2 a - (|a + F|^2(a + F) - |a|^2 a). \tag{3.26}$$

In the rest of the article we will need two parameters s, δ. We always assume

$$0 < \delta \ll s \ll 1, \tag{3.27}$$

and in particular for any small constant c involved in our analysis, one has

$$(1 - cs)(1 + \delta) < 1. \tag{3.28}$$

Now we are ready to state the main bootstrap lemma. Let u solve (1.1) in $[0, T]$ with initial data u_0 described as in §3.1, with ansatz (3.20) so that (3.21)–(3.24) hold. Since $\lambda(t)$ is essentially decreasing, or more precisely by the bootstrap assumption (3.30), we can divide $[0, T]$ into $\bigcup_{k=k_0}^{k_1=1}[t_k, t_{k+1}]$ so that $\lambda(0) \sim 2^{-k_0}$ and $\lambda(T) \sim 2^{-k_1}$, $k_0 \le k \le k_1$, and $\lambda(t) \sim 2^{-k}$ for all $t \in [t_k, t_{k+1}]$.

Lemma 3.6 (Bootstrap lemma). *Suppose that $u(t, x)$ solves (1.1) on $[0, T]$ and satisfies the following bootstrap assumptions for $t \in [0, T]$:*

$$0 < b(t), \quad \|\epsilon\|_{L^2} + b(t) < \alpha, \tag{3.29}$$
$$\forall t \le t' \in [0, T], \quad \lambda(t') \le \tfrac{3}{2}\lambda(t), \tag{3.30}$$
$$\lambda(t) \le e^{-\Gamma_b^{-2/3}}, \tag{3.31}$$
$$\int |\nabla I_{N(t)\lambda(t)}\epsilon|^2 + |\epsilon|^2 e^{-|y|} \le \Gamma_b^{2/3}, \tag{3.32}$$
$$t_{k+1} - t_k \lesssim k\lambda(t_k)^2 \sim k2^{-2k}. \tag{3.33}$$

Then

$$0 < b(t), \quad \|\epsilon\|_{L^2} + b(t) < \alpha/2, \tag{3.34}$$
$$\forall t \le t' \in [0, T], \quad \lambda(t') \le \tfrac{5}{4}\lambda(t), \tag{3.35}$$
$$\lambda(t) \le e^{-\Gamma_b^{-3/4}}, \tag{3.36}$$
$$\int |\nabla I_{N(t)\lambda(t)}\epsilon|^2 + |\epsilon|^2 e^{-|y|} \le \Gamma_b^{3/4}, \tag{3.37}$$
$$t_{k+1} - t_k \lesssim \sqrt{k}\,\lambda(t_k)^2 \sim \sqrt{k}\,2^{-2k}, \tag{3.38}$$

Remark 3.7. Formally, the asymptotic dynamics gives

$$b_s \sim -\Gamma_b, \quad \int |I_{N(t)\lambda(t)}\epsilon|^2 + |\epsilon|^2 \lesssim \Gamma_b,$$
$$\lambda \sim e^{-\Gamma_b^{-1}}, \quad t_{k+1} - t_k \sim \log\log k,$$

and the mass conservation law gives

$$\|\epsilon\|_2 + b^2 \lesssim \|a(0)\|_{L_x^2} - \|Q_0\|_{L_x^2}.$$

Also the condition $b_s \sim -\Gamma_b$ essentially ensures that b stays positive for all time.

Remark 3.8. Note that if $f_\lambda = \frac{1}{\lambda} f(x/\lambda)$, then $I_N f_\lambda = (I_{N\lambda} f)_\lambda$. We will use this repeatedly when computing quantities from the bootstrap lemma.

4. Probabilistic local well-posedness

We note that while the cubic nonlinear Schrödinger equation (1.1) is *deterministically* well-posed in $L_x^2(\mathbb{R}^2)$, we are seeking nonlinear smoothing and quantitative estimates which are not true for general deterministic data. Hence, we exploit several properties of the free evolution of random data, as well as multilinear estimates involving such random functions.

The analysis in this section has many similarities to the random data analysis of Bourgain [5]. Indeed, our choice of function to randomize is intended to mimic the random data appearing in [5]. However, several new ingredients are needed to carry out these estimates, and in particular, we need some new arguments in order to adapt Bourgain's result to the noncompact setting.

For technical reasons, we fix $\varepsilon_0 > 0$ and $b = 1/2 + \epsilon_0$. We will also fix

$$\epsilon_0 \ll \epsilon_1 \ll \epsilon_2 \ll \delta \ll s \ll 1, \tag{4.1}$$

and any ϵ involved in the analysis should satisfy $\epsilon \ll \epsilon_0$.

Remark 4.1. One may assume, for example, $\epsilon_0 \ll \epsilon_2^{s/10}$. The purpose of these parameters is to overcome a technical issue arising from the scaling of $X^{s,b}$ spaces, specifically letting $h_\lambda := \frac{1}{\lambda} h(t/\lambda^2, x/\lambda)$, and $b = 1/2 + \epsilon_0$, one has

$$\|h_\lambda\|_{X^{s,b}} \lesssim \|h\|_{X^{s,b}} \lambda^{-s-2\epsilon_0}.$$

We note that in general, the scaling properties of $X^{s,b}$ do not pose problems since our local well-posedness and energy estimates will be subcritical in nature, and we do not need to derive endpoint type estimates where ϵ losses would be forbidden.

The aim of the current section is to establish improved[18] probabilistic local well-posedness. Specifically, we will establish a result analogous to [20, Lemma 3.3] with randomized data.

In [20], every LWP interval $[t_k, t_{k+1}]$ is split into $\bigcup_j [\tau_k^j, \tau_k^{j+1}]$ so that

$$|\tau_k^j - \tau_k^{j+1}| \sim \lambda(\tau_k^j)^{-2} \sim \lambda(t_k)^{-2} \sim 2^{-2k}.$$

Due to the aforementioned technical issues relating to the scaling of $X^{s,b}$ spaces, we will instead split $[t_k, t_{k+1}]$ into $\bigcup_{j=1}^{J_k} [\tau_k^j, \tau_k^{j+1}]$ so that

$$|\tau_k^j - \tau_k^{j+1}| \sim \lambda(t_k)^{2-\epsilon_2} \tag{4.2}$$

and note there are at most $k\lambda(t_k)^{-\epsilon_2}$ such LWP intervals within $[t_k, t_{k+1}]$, thanks to the bootstrap assumption (3.33).

[18] It is standard that the problem we treat in this article is deterministically locally well-posed with intervals of length $\sim \lambda(t_k)^{-2+\epsilon_2}$ if one only cares about L^2 level well-posedness.

Thus, let $0 < T \ll 1$ and let $I = [\tau_k^j, \tau_k^{j+1}] \subset [t_k, t_{k+1}] \subset [0, T]$ with $|I| \sim \lambda(t_k)^{2-\epsilon_2}$. Recall u solves

$$\begin{cases} iu_t + \Delta u = -|u|^2 u, & (x,t) \in \mathbb{R}^2 \times I, \\ u(\tau_k^j) = a(\tau_k^j) + F(\tau_k^j), \end{cases} \tag{4.3}$$

where a is of the form (3.20) and satisfies the bootstrap assumption (3.32). We note that (3.32) implies

$$\|a(\tau_k^j)\|_{H^s} \sim \frac{1}{\lambda(\tau_k^j)^s} \sim \frac{1}{\lambda(t_k)^s}. \tag{4.4}$$

We now turn to probabilistic local estimates.

Lemma 4.2. *Let f^ω be the randomization defined in (3.1). Fix $p = \infty-$ and $q = 4$, and let $\Sigma_1 \subseteq \Omega$ be a subset such that (2.5), (2.12) and (2.13) hold. Then there exists a set $\Sigma_2 \subseteq \Omega$ satisfying*

$$\mathbb{P}(\Sigma_2^c) \lesssim e^{-|\tau_k^j - \tau_k^{j+1}|^{-c}}$$

such that for every $\omega \in \Sigma_1 \cap \Sigma_2$, if u solves (1.1) with initial data $a_0 + f^\omega$, then

$$\|a\|_{X^{s,b}[I]} = \|u - F\|_{X^{s,b}[I]} \lesssim \frac{1}{\lambda(\tau_k)^{s+\epsilon_1}}, \tag{4.5}$$

and

$$\|I_{N(\tau_k^j)} a\|_{X^{1,b}[I]} = \|I_{N(\tau_k^j)}(u - F)\|_{X^{1,b}[I]} \sim \frac{1}{\lambda(t_k)^{1+\epsilon_1}}. \tag{4.6}$$

Note that in particular (4.6) implies

$$\|I_{N(T)} a\|_{X^{1,b}[I]} \lesssim \left(\frac{N(T)}{N(t_k)}\right)^{1-s} \left(\frac{1}{\lambda(t_k)}\right)^{1+\epsilon_1} \tag{4.7}$$

(see [20, (3.14) and (3.20)]).

Remark 4.3. We may establish an identical result and additionally

$$\mathbb{P}(\Sigma_1^c) \lesssim e^{-1/\lambda_0^c}$$

for some $c > 0$. Indeed, fix $p = \infty-$, and let $\Sigma_1 \subseteq \Omega$ be such that (2.5), (2.12) and (2.13) hold with constant $\sim \lambda_0^{-c_1}$ for some $c_1 > 0$ small. Then up to redefining ε_1, this additional loss can be absorbed into estimates (4.5) and (4.6). We additionally note that such a subset is independent of k and $I = [\tau_k^j, \tau_k^{j+1}]$. Hence, to simplify our arguments, we will instead assume that (2.5), (2.12) and (2.13) hold with a fixed (λ_0-independent) constant.

Remark 4.4. On the whole interval $[0, T]$, the set that one needs to drop arising from the subset Σ_2 in Lemma 4.2 contributes total probability[19] bounded by

$$\sum_{k=k_0}^{k_+} k e^{-2^{ck}} \leq \sum_{k=k_0}^{k_+} k e^{-2^{ck}} \lesssim C e^{-2^{c/2} k_0}. \tag{4.8}$$

[19] The c may change from line to line and may not be the same as in Lemma 4.2.

Thus, by making k_0 large enough (i.e. λ_0 small enough), one can ensure that up to a set of small probability, for every $I = [\tau_k^j, \tau_k^{j+1}] \subset [t_k, t_{k+1}] \subset [0, T]$, the conclusion of Lemma 4.2 holds.

Remark 4.5. As we will see, the proof reduces to controlling the nonlinear expression $|a + F|^2(a + F)$, and in particular the term $|F|^2 F$ is the most difficult to control.

While the term $|a|^2 a$ essentially follows from standard deterministic theory, we need to introduce parameters $\epsilon_0, \epsilon_1, \epsilon_2$ in (4.1) for the following reasons: we will need to rescale a to $\lambda a(\lambda^2 t, \lambda x)$ so that it is normalized in $X^{s,b}$, apply the standard deterministic local theory, and then scale back. This generates an extra error $\lambda(t_k)^{-C\epsilon_0}$ due to the fact that $X^{s,b}$ is not scale invariant, resulting in an extra loss of $\lambda(t_k)^{-\epsilon_1}$.

The term $|a|^2 F$ also essentially follows from deterministic local theory since we are able to distribute derivatives using bilinear Strichartz estimates, and have sufficiently smooth functions to do so. However, we need the smallness of the interval to close this estimate, and hence we shrink the interval by an extra $\lambda(t_k)^{\epsilon_2}$ factor. We will not focus too strongly on these parameters since we wish to emphasize the treatment of the terms $|F|^2 F$ and $|F|^2 a$ (up to complex conjugates); however, we point out that any loss of the form of $\lambda(t_k)^{-\epsilon_2}$ is acceptable in the estimates because of the smallness of the time interval.

We note in particular that the gain due to the smallness of the interval does not follow from extra room in the b parameter of $X^{s,b}$ spaces.

Proof of Lemma 4.2. Recall that we use the ansatz (3.20), and that a solves the difference equation

$$\begin{cases} ia_t + \Delta a = |a + F|^2(a + F), & t \in [\tau_k^j, \tau_k^{j+1}], \\ \|a(\tau_k, x)\|_{H^s} \sim \lambda(\tau_k)^{-s}. \end{cases} \tag{4.9}$$

Without loss of generality and by a time translation, we may take $\tau_k^j = 0$. Let $\eta(t)$ be a smooth cut-off, with $\eta(t) \equiv 1$ when $|t| \leq 1$, and $\eta(t) \equiv 0$ for $|t| \geq 2$. Let $\eta_\beta(t) = \eta(t/\beta)$. We denote by $\widetilde{a}$ the extension of a over the real line. By Duhamel's formulation, we need to estimate

$$a(t, x) = e^{i(t-\tau_k)} a(\tau_k, x) - i \int_{\tau_k}^t e^{-i(t-s)} \big(|a + F|^2(a + F)\big)\, ds.$$

The linear part of $a(t, x)$ can be handled with the standard $X^{s,b}$ estimate, using the form of $a(\tau_k^j)$. For the inhomogeneous nonlinear estimate we need to control

$$\begin{aligned} &\left\| \int_{\tau_k}^t e^{-i(t-s)} \big(|a + F|^2(a + F)\big) ds \right\|_{X^{s,b}(I)} \\ &\quad \lesssim \left\| \eta_{|I|} \int_{\tau_k}^t e^{-i(t-s)} \big(|\widetilde{a} + F|^2(\widetilde{a} + F)\big) ds \right\|_{X^{s,b}} \\ &\quad \lesssim \left\| \eta_I(t) |\widetilde{a} + F|^2(\widetilde{a} + F)| \right\|_{X^{s,b-1}} \\ &\quad \lesssim \left\| \eta_I(t) \big(|\widetilde{a} + F|^2(\widetilde{a} + F) - |F|^2 F\big) \right\|_{X^{s,b-1}} + \left\| \eta_I(t) |F|^2 F \right\|_{X^{s,b-1}} \end{aligned} \tag{4.10}$$

Our main goal is to prove that given $\widetilde{a}$ which satisfies

$$\|\widetilde{a}\|_{X^{s,b}} \lesssim \left(\frac{1}{\lambda(\tau_k^j)}\right)^s,$$

(4.10) is bounded by

$$\frac{1}{2}\left(\frac{1}{\lambda(\tau_k^j)}\right)^s. \tag{4.11}$$

The extra $1/2$ factor implies, in the usual manner, that the solution map is a contraction. Indeed, this follows from the fact that the $X^{s,b}$ and random data analysis involved is subcritical in nature, and that we are working on a small interval. We note that in order to establish (4.11), it will suffice to prove that (4.10) is bounded by

$$\left(\frac{1}{\lambda(\tau_k^j)}\right)^{s+\epsilon_1} \tag{4.12}$$

since we are working on intervals of (extra) small length, and the extra smallness of the time interval, $\lambda(t_k)^{\epsilon_2}$, will be able to beat the ϵ_1 loss, as remarked in Remark 4.5. Below, we will not distinguish between a and $\widetilde{a}$, since they will be treated and estimated in the same way. We finally remark that there is a simple way to gain smallness of $X^{s,b}$ by localizing time, i.e. to use part of b derivatives to estimate $X^{s,b'}$ for some $1/2 < b' < b$. This will *never* be involved in our analysis, however, because the maximum allowable difference between b and b' is bounded by ϵ_0, which is too small to overcome the extra loss in (4.12).

Thus, we focus on establishing (4.12), and we begin with the term $\eta_I(t)|F|^2F$. In light of our discussion in Remark 4.3, we will prove that

$$\left\|\eta_{|I|}(t)|F|^2F\right\|_{X^{s,b-1}} \lesssim 1. \tag{4.13}$$

Moreover, for the majority of the proof, we will in fact prove that

$$\left\|\eta(t)|F|^2F\right\|_{X^{s,b-1}} \lesssim 1, \tag{4.14}$$

and we note that we can replace the term $\eta_{|I|}(t)$ with $\eta(t)$ since $X^{s,b}$ spaces are well-behaved under time localization. Additionally, we will occasionally abuse notation, and use $\eta^3 \simeq \eta$, which will enable us to replace F with $\eta(t)F$ as needed. Time localization is only needed when we argue that the extra subset we drop has probability $\lesssim e^{-|\tau-\tau_{j+1}|^{-c}}$. We will revisit this later.

Let

$$* = \{n_1, n_2, n_3 \in \mathbb{Z}^2,\ n_2 \neq n_1, n_3\}$$

and set

$$h_k = e^{it\Delta}\check{\psi}_k.$$

Then $|F|^2F$ can be written as

$$\underbrace{\sum_{*}\iiint g_{n_1}\overline{g}_{n_2}g_{n_3}f_{n_1}\overline{f}_{n_2}f_{n_3}\psi_{n_1}(\xi_1)\overline{\psi}_{n_2}(\xi_2)\psi_{n_3}(\xi_3)e^{i(\xi_1-\xi_2+\xi_3)x-(|\xi_1|^2-|\xi_2|^2+|\xi_3|^2)t}}_{\text{Term 1}}$$

$$-\underbrace{\sum_{n}\iiint |g_n|^2g_n|f_n|^2f_n\psi_n(\xi_1)\overline{\psi}_n(\xi_2)\psi_n(\xi_3)e^{i(\xi_1-\xi_2+\xi_3)x-(|\xi_1|^2-|\xi_2|^2+|\xi_3|^2)t}}_{\text{Term 2}}$$

$$+\underbrace{2\sum_{n_1,n_3}|f_{n_1}|^2|g_{n_1}|^2|h_{n_1}|^2f_{n_3}g_{n_3}h_{n_3}}_{\text{Term 3}}.$$

We estimate these terms separately, beginning with the easiest.

Term 2. As in [5] we will directly estimate the $L_t^\infty H_x^s$ norm of this term. A simple triangle inequality gives

$$\Big\|\langle\xi\rangle^s\sum_n |g_n|^2g_n|f_n|^2f_n\iiint \delta(\xi-\xi_1+\xi_2-\xi_3)\psi_k(\xi_1)\psi_k(\xi_2)\psi_k(\xi_3)\Big\|_{L^2_\xi}$$
$$\lesssim \sum_n\langle n\rangle^s\langle n\rangle^{3\varepsilon}|n|^{-3}\lesssim 1 \qquad (4.15)$$

where we have used the ℓ^∞ Gaussian bound (2.5). This is summable in dimension $d=2$ provided $s<1$.

Term 1. This is the term which typically appears in random data analysis, and can usually be used to illustrate what improvements one obtains for random data; see [5] for more details. Here, since we are not working with the NLS on a (rational) torus, one cannot directly reduce the problem to the same counting estimates as in [5]. On the other hand, since we are on Euclidean space, we can take advantage of the bilinear Strichartz estimates in Lemma 2.5.

Below we assume $|\xi_i|=|k_i|+O(1)\sim N_i$ for $i=1,2,3$ and without loss of generality, set $N_1\geq N_2\geq N_3$, where N_i are dyadic integers. We write $F_i:=F_{N_i}:=P_{N_i}F$, $i=1,2,3$. We first perform several reductions. We note that we may assume that

$$N_3\geq N_1^{99/100}, \qquad (4.16)$$

since otherwise by bilinear Strichartz estimates, we obtain (recall p is always large)

$$\begin{aligned}\|\eta(t)F_1F_2F_3\|_{X^{s,b-1}} &\lesssim \sup_{\|h\|_{X^{s,1-b}}=1}\int F_1F_2F_3\bar{h}\\ &\lesssim N_1^s\|F_1F_3\|_{L_t^2L_x^2}\|F_2\|_{L^p_{t,x}}\\ &\leq N_1^s\left(\frac{N_3}{N_1}\right)^{1/2}\|F_2\|_{L^p_{t,x}},\end{aligned} \qquad (4.17)$$

and using (2.12), we may sum ($N_1 \geq N_2 \geq N_3$) provided

$$N_1^{s-\frac{1}{2}+\frac{2}{p}} N_3^{\frac{1}{2}} \leq N_1^{s-\frac{1}{2}+\frac{198}{100p}} N_1^{\frac{99}{200}} \lesssim 1,$$

which can be done by choosing $s > 0$ sufficiently small so that

$$s - \frac{1}{200} + \frac{198}{100p} < 0.$$

Proceeding, we will estimate this expression by reducing the problem to counting problems. Following Bourgain [5], we start with a standard reduction. Here, we need to replace F by $\eta(t)F$. By definition of the $X^{s,b}$ space, we need to control

$$\|\langle \tau - |\xi|^2 \rangle^{b-1} \langle \xi \rangle^s \mathcal{F}_{\tau,\xi}(\eta(t)\text{Term } 1)\|_{L^2_\tau L^2_\xi}, \tag{4.18}$$

We let $\mu = \tau - |\xi|^2$ and we first claim we only need to control the region

$$\mu \ll N_1^{10s}. \tag{4.19}$$

Indeed, one may use dual estimates to estimate (4.18). For the deterministic theory, one needs to pair a function h such that $\|h\|_{X^{0,1-b}} = 1$, and all $1 - b = 1/2 + \epsilon_0$ ($X^{s,b}$ type) derivatives are needed, since one needs to control $\|h\|_{L^4_{t,x}}$. Here, the random data allows us to beat the usual Strichartz estimates, and we are able to place each copy of F in $L^{\infty-}_{t,x}$, and hence one only needs control of $\|h\|_{L^{3+}_{t,x}}$, which by interpolation only requires $1/3 + \epsilon$ ($X^{s,b}$ type) derivatives. Thus, if one is in the case $\|\mu\| \geq N_1^{10s}$, the gain in the $X^{s,b}$ smoothing will compensate the N_1^s loss in the space derivative; see [5, (30) and (35)]. Moreover, we may focus on the case $\mu = O(1)$ and sum different parts via the triangle inequality, suffering an extra N_1^{Cs} loss; note C will be large but we still have $Cs \ll 1$.

Going back to (4.18), we first expand $\mathcal{F}(\eta(t)\text{Term } 1)$:

$$\begin{aligned}&\iint e^{-ix\cdot\xi} e^{-i\tau t} \eta(t) \iiint \psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_3) e^{i(\xi_1-\xi_2+\xi_3)x-(|\xi_1|^2-|\xi_2|^2+|\xi_3|^2)t} \\ &\quad = \iiint \hat{\eta}(\tau - |\xi_1|^2 - |\xi_2|^2 + |\xi_3|^2)\delta(\xi - \xi_1 + \xi_2 - \xi_3)\psi_{n_1}(\xi_1)\psi_{n_2}(\xi_2)\psi_{n_3}(\xi_3).\end{aligned} \tag{4.20}$$

We substitute this expression into (4.18), and we recall that

$$|\xi_1 - \xi_2 + \xi_3|^2 - |\xi_1|^2 + |\xi_2|^2 - |\xi_3|^2 = 2\langle \xi_2 - \xi_1, \xi_2 - \xi_3 \rangle.$$

Now, we need to argue separately for $N_1 \geq N_{1,0}$ and $N_1 < N_{1,0}$ for some $N_{1,0}$ which we will determine below. When $N_1 \geq N_{1,0}$, this is where we drop the extra set of small probability, Σ_2^c, mentioned in the statement of the lemma. This extra argument is (more or less) standard, but we provide a sketch here. We fix such an N_1, and we use the multilinear Gaussian estimate of Lemma 2.14 with constant $K = N_1^{Cs}$ to replace Term 1 by its L^2_ω

norm by dropping an *extra* set of probability $\leq e^{-N_1^{c(\epsilon)}}$, where $c(\epsilon) > 0$ is a small ϵ-dependent constant. Ultimately we need to control[20]

$$N_1^{2Cs}\Big(\sum_{*}\frac{1}{|n_1|^2}\frac{1}{|n_2|^2}\frac{1}{|n_3|^2}\Big)^{1/2} \tag{4.21}$$

where

$$* = \{n_1, n_2, n_3,\ n_2 \neq n_1, n_3,\ \langle n_2 - n_3, n_2 - n_1\rangle = O(N_1),\ |n_i| \sim N_i\}.$$

Recall that by restricting to the case $\mu = O(1)$ we lose an extra N_1^{Cs}, and we a priori have $|f_{n_i}| \leq 1/|n_i|$. As in (4.16), we only consider the case $N_3 \geq N_1^{99/100}$.

When $|N_2 - N_3| < N_3^{1/10}$, for fixed n_1, n_2 there will be at most $N_3^{1/5}$ many n_3, and we may use $N_3 \geq N_1^{99/100}$ to sum

$$\sum_{N_1, N_2, N_3, |N_2 - N_3| < N_3^{1/10}} N_1^{2Cs}\frac{1}{|n_1|^2}\frac{1}{|n_2|^2}\frac{1}{|n_3|^2} \lesssim N_1^{2Cs} N_1^{-(2-\frac{1}{5})\cdot\frac{99}{100}},$$

which is acceptable for s sufficiently small.

When $|N_2 - N_3| \geq N_3^{1/10}$, we mimic the counting in [5, Lemma 1]. Fixing n_2 and n_3, we note there could be at most $N_1^2/N_3^{1/10}$ many n_1. Indeed, let

$$n_1 - n_3 = (c_1, c_2), \quad n_2 - n_3 = (b_1, b_2),$$

and assuming, for example, that $b_2 \geq N_3^{1/10}$, and fixing c_1, there can be at most $N_1/N_3^{1/10}$ many c_2, and at most N_1 many a_1. Hence, we may bound (4.21) by $N_3^{-1/10}$ or $N_1^{-1/20}$, since $N_3 \geq N_1^{99/100}$, and we obtain a bound which is summable for s sufficiently small. Since we must drop an extra subset for every fixed $N_1 \geq N_{1,0}$, after summation we see that the probability of the subset we drop is $\leq e^{-(N_{1,0})^c}$.

For $N_1 \leq N_{1,0}$, we use (2.5), and then we argue in a purely deterministic manner, using the fact the interval is (extra) short, of length $\sim \lambda(t_k)^{-2+\epsilon_2}$ to close. Here, we need to use the cut-off $\eta_{|I|}(t)$. To close these estimates, we fix $N_{1,0} \sim \lambda(t_k)^{-\tilde{c}(\epsilon)}$, where again $\tilde{c}(\epsilon) > 0$ is another small, ϵ-dependent constant. This yields the stated bound on $\mathbb{P}(\Sigma_2^c)$, recalling how we defined the length of the time intervals. See also the discussion in [5, below (46)].

Term 3. This term is the most distinct from the analysis in [5]. Indeed, in [5], a Wick ordering is applied and this term does not appear at all. We remark that one can still apply a phase transform to cancel this term, but such a phase, unlike one in [5], will be a function

[20] Here the situation is different from the tori case in [5]. In the periodic case, the constraint $\mu = O(1)$, or $\langle \xi_2 - \xi_1, \xi_2 - \xi_3\rangle = O(1)$, is reduced to exact $\langle n_2 - n_3, n_2 - n_1\rangle = O(1)$. However, here, we only have $|\xi_i - n_i| \sim 1$, thus we only get $\langle n_2 - n_3, n_2 - n_1\rangle = O(N_1)$ and extra effort will be needed to control the summation.

rather than a number, and will not leave the NLS invariant. The key difference between our setting and Bourgain's is that our initial data lies at $L^2_x(\mathbb{R}^2)$ regularity, and hence we do not have to control the same divergences which appear for data in the support of the invariant Gibbs measure considered by Bourgain.

We recall that we are considering the term

$$\sum_{n_1,n_3} |f_{n_1}|^2 |g_{n_1}|^2 |h_{n_1}|^2 f_{n_3} g_{n_3} h_{n_3};$$

we let

$$\widetilde{\theta}(t,x,\omega) = 2\sum_{n_1} |f_{n_1}|^2 |g_{n_1}|^2 |h_{n_1}|^2,$$

and note that this term is equal to $\widetilde{\theta}(t,x,\omega)F$. Moreover, we observe that

$$\mathbb{E}\Big(\sum_{n_1} |f_{n_1}|^2 |g_{n_1}|^2\Big) < \infty,$$

and hence almost surely $\{f_{n_1} g_{n_1}\}_{n_1\in\mathbb{Z}^2} \in \ell^2$ and up to an exceptional set from (2.5), we have

$$|f_{n_1} g_{n_1}| \lesssim \frac{|n_1|^{\varepsilon}}{|n_1|}.$$

Now, observing that

$$|h_k|^2 = e^{it\Delta}\check{\psi}_k\, \overline{e^{it\Delta}\check{\psi}_k},$$

and using the fact that the free evolution does not affect the Fourier support, for each k the term $|h_k|^2$ is supported in a ball of radius 2 around the origin by convolution of the supports. And indeed $|h_k|^2 = |h_0(t, x-kt)|^2$ and h_0 is smooth. Thus, $\theta(x,t)$ is also smooth since $\sum |f_k|^2 \lesssim 1$.

We need to estimate $\|\widetilde{\theta}(t,x,\omega)F\|_{L^2_t H^s_x}$, and in light of the observations above, it suffices to estimate the expression

$$\big\|\widetilde{\theta}(t,x,\omega)|\nabla|^s F\big\|_{L^2_t L^2_x}.$$

First observe that since $|h_k|^2$ are all frequency localized around 1, we apply L^2-orthogonality to derive

$$\big\|\widetilde{\theta}(t,x,\omega)|\nabla|^s F\big\|_{L^2_t L^2_x} \lesssim \sum_{k\in\mathbb{Z}^2} \big\|\widetilde{\theta}(t,x)|\nabla|^s P_k f\big\|^2_{L^2_t L^2_x}. \tag{4.22}$$

Now, noting that h_ℓ enjoys unit-scale Bernstein estimates (and hence lies in $L^\infty_{t,x}$) we obtain

$$\begin{aligned}\sum_{k\in\mathbb{Z}^2} &\|\widetilde{\theta}(t,x)e^{it\Delta}P_k f\|^2_{L^2_t L^2_x}\\ &\lesssim \sum_k \Big(\sum_\ell |f_\ell g_\ell|^2 \|h_\ell\|_{L^\infty_{t,x}} \big\|e^{it\Delta}\check{\psi}_\ell\, e^{it\Delta}(g_k\psi_k)(\omega)|k|^s f_k\big\|_{L^2_t L^2_x}\Big)^2. \end{aligned}\tag{4.23}$$

Now, we apply the bilinear Strichartz estimate of Proposition 2.5, and conjugating with Galilean symmetry for

$$e^{it\Delta}\check{\psi}_\ell\, e^{it\Delta}(g_k\psi_k)(\omega),$$

and plugging in the ℓ^∞ Gaussian bound (2.5), we can estimate this expression by

$$\sum_k |f_k|^2\Big(\sum_\ell |f_\ell|^2|k|^{s+\varepsilon}\frac{1}{\langle k-\ell\rangle^{1/2}}\Big)^2. \tag{4.24}$$

Note that $|f_k|^2$ is summable in k.

Now, fix k. If $|k-\ell| > k/2$, then provided $s+\varepsilon < 1/2$, this expression is bounded. Alternatively, if $|k-\ell| \le k/2$, then we use the fact that $\||\ell|f_\ell\|_{\ell^\infty} \le C$ and $\ell \sim k$ to obtain

$$\begin{aligned}\sum_{\ell,\,|\ell-k|\le k/2} &|f_\ell|^2|k|^{s+\varepsilon}\frac{1}{\langle k-\ell\rangle^{1/2}} \\ &\lesssim \sum_{\ell,\,|\ell-k|\le k/2} k^{-2}|k|^{s+\varepsilon}\frac{1}{\langle k-\ell\rangle^{1/2}} \simeq k^{-2}|k|^{s+\varepsilon}|k|^{3/2},\end{aligned} \tag{4.25}$$

which is bounded provided, again, $s+\varepsilon < 1/2$.

We now turn to the terms involving $\widetilde{a}$. For notational convenience, we will still use a to denote $\widetilde{a}$. Again, we write $a_i := a_{N_i} := P_{N_i}a$, and N_i is a dyadic integer, and similarly for F_j, $j = 1,2,3$. We will have to deal with multiple cases, depending on the frequency at which the random function is appearing. Before proceeding, we note that since we are working on an interval of length $\lesssim \lambda(t_k)^2$, and by assumptions on the subset of the probability space, we can use Hölder's inequality in time to break the scaling and derive, for example,

$$\|F\|_{L^4_{t,x}[I]} \lesssim |I|^{\frac{1}{4}-} \lesssim |\lambda(t_k)|^{\frac{1}{2}-}. \tag{4.26}$$

This will be frequently used in the analysis below.

Case 1: $\|F_{N_1}F_{N_2}a_{N_3}\|$. By duality, we estimate the expression

$$\sum_{N_1\gg N_2\ge N_3} N_1^s\int F_1F_2a_3h$$

for $h \in X^{0,1-b}$. When $N_1 \sim N_2$, we estimate this expression using bilinear Strichartz estimates and Cauchy–Schwarz in the highest frequency:

$$\begin{aligned}&\sum_{N_1\sim N_2\ge N_3}\Big(\frac{N_3}{N_1}\Big)^{1/2-s}\|F_1\|_{L^\infty_t L^2_x}\|F_2\|_{L^{4+}_{t,x}}N_3^s\|a_3\|_{X^{0,b}}\|h\|_{L^{4-}_{t,x}} \\ &\lesssim \sum_{N_1\sim N_2\ge N_3}\Big(\frac{N_3}{N_1}\Big)^{1/2-s}\|F_1\|_{L^\infty_t L^2_x}\lambda(t_k)^{\frac{1}{2}-}\|F_2\|_{L^{\infty-}_t L^{4+}_x}N_3^s\|a_3\|_{X^{0,b}}\|h\|_{L^{4-}_{t,x}},\end{aligned} \tag{4.27}$$

which is summable.

When $N_1 \gg N_2$, we use duality with $h \in X^{0,1-b}$, and we decompose h into dyadic blocks h_{N_4}, now having $N_4 \sim N_1$, and once again by bilinear Strichartz estimates,

$$\begin{aligned}
&\sum_{N_1 \gg N_2 \geq N_3} N_1^s \int F_1 F_2 a_3 h \\
&\lesssim \sum_{N_4 \sim N_1 \gg N_2 \geq N_3} \left(\frac{N_3}{N_1}\right)^{1/2-s} \|F_1\|_{L_t^\infty L_x^2} \|h_4\|_{X^{0,1-b}} \|F_2\|_{L_{t,x}^{4+}} N_3^s \|a_3\|_{X^{0,b}} \\
&\lesssim \sum_{N_4 \sim N_1 \gg N_2 \geq N_3} \left(\frac{N_3}{N_1}\right)^{1/2-s} \|F_1\|_{L_t^\infty L_x^2} \|h_4\|_{X^{0,1-b}} \lambda(t_k)^{\frac{1}{2}-} \|F_2\|_{L_t^{\infty-} L_x^{4+}} N_3^s \|a_3\|_{X^{0,b}}.
\end{aligned} \tag{4.28}$$

This is again summable.

Case 2: $\|a_{N_1} F_{N_2} F_{N_3}\|$. This can be estimated precisely as in the previous estimate, but we do not need to transfer regularity through bilinear Strichartz.

Case 3: $\|F_{N_1} a_{N_2} F_{N_3}\|$. When $N_1 \sim N_2$, We estimate using duality and bilinear Strichartz:

$$\begin{aligned}
&\sum_{N_1 \sim N_2 \geq N_3} N_1^s \int F_1 a_2 F_3 h \\
&\quad \lesssim \sum_{N_1 \sim N_2 \geq N_3} N_1^s \|F_1 h\|_{L_{t,x}^2} \|a_2 F_3\|_{L_{t,x}^2} \\
&\quad \lesssim \sum_{N_1 \sim N_2 \geq N_3} N_1^s N_2^{-s} \|F_1\|_{L_{t,x}^{4+}} \|a_2\|_{X^{s,b}} \|F_3\|_{L_t^\infty L_x^2} \left(\frac{N_3}{N_2}\right)^{1/2} \\
&\quad \simeq \sum_{N_1 \sim N_2 \geq N_3} N_1^s N_2^{-s} \|F_1\|_{L_{t,x}^{4+}} \|a_2\|_{X^{s,b}} \|F_3\|_{L_t^\infty L_x^2} \left(\frac{N_3}{N_1}\right)^{1/2} \\
&\quad \lesssim \sum_{N_1 \sim N_2 \geq N_3} N_1^s N_2^{-s} \lambda(t_k)^{\frac{1}{2}-} \|F_1\|_{L_t^{\infty-} L_x^{4+}} \|a_2\|_{X^{s,b}} \|F_3\|_{L_t^\infty L_x^2} \left(\frac{N_3}{N_1}\right)^{1/2},
\end{aligned} \tag{4.29}$$

and we can sum this expression.

When $N_1 \gg N_2$, we use duality with $h \in X^{0,1-b}$, and we decompose h into dyadic blocks h_{N_4}, now having $N_4 \sim N_1$. We estimate

$$\sum_{N_4 \sim N_1 \gg N_2 \geq N_3} N_1^s \int F_1 a_2 F_3 h_4. \tag{4.30}$$

We pair a_2 with either F_1 or F_3 depending on the value of

$$\min\left(\left(\frac{N_2}{N_1}\right), \left(\frac{N_3}{N_2}\right)\right),$$

using the other F factor to estimate with h as above.

For example, supposing we perform the bilinear Strichartz with $a_2 F_3$, (in the case $N_3/N_2 \le N_2/N_1$) we then obtain

$$\begin{aligned}
&\sum_{N_4 \sim N_1 \gg N_2 \ge N_3} N_1^s N_2^{-s} \left(\frac{N_3}{N_2}\right)^{1/2} \|F_3\|_{L_t^\infty L_x^2} \|a_2\|_{X^{s,b}} \|h_4\|_{X^{0,1-b}} \|F_1\|_{L_{t,x}^{4+}} \\
&= \sum_{N_4 \sim N_1 \gg N_2 \ge N_3} N_1^s N_2^{-s} \left(\frac{N_3}{N_1}\right)^{1/4} \|F_3\|_{L_t^\infty L_x^2} \|a_2\|_{X^{s,b}} \|h_4\|_{X^{0,1-b}} \|F_1\|_{L_{t,x}^{4+}} \\
&\lesssim \sum_{N_4 \sim N_1 \gg N_2 \ge N_3} N_1^s N_2^{-s} \left(\frac{N_3}{N_1}\right)^{1/4} \|F_3\|_{L_t^\infty L_x^2} \|a_2\|_{X^{s,b}} \|h_4\|_{X^{0,1-b}} \lambda(t_k)^{\frac{1}{2}-} \|F_1\|_{L_t^{\infty-} L_x^{4+}}
\end{aligned} \tag{4.31}$$

where we have used $\min(a,b) \le \sqrt{ab}$. Once again this is summable for $s < 1/4$ using Cauchy–Schwarz in $N_1 \sim N_4$.

Case 4: $\|F_{N_1} a_{N_2} a_{N_3}\|$. Once again, we estimate by duality. If $N_1 \sim N_2$, we have

$$\sum_{N_1 \gg N_2 \ge N_3} N_1^s \int F_1 a_2 a_3 h$$

and we estimate using bilinear Strichartz with $a_2 a_3$:

$$\begin{aligned}
&\sum_{N_1 \gg N_2 \ge N_3} N_1^s N_2^{-s} N_3^{-s} \left(\frac{N_3}{N_2}\right)^{1/2} \|F_1\|_{L_{t,x}^{4+}} \|a_2\|_{X^{s,b}} \|a_3\|_{X^{s,b}} \|h\|_{X^{0,1-b}} \\
&\lesssim \sum_{N_1 \gg N_2 \ge N_3} N_1^s N_2^{-s} N_3^{-s} \left(\frac{N_3}{N_2}\right)^{1/2} \lambda(t_k)^{\frac{1}{2}-} \|F_1\|_{L_t^{\infty-} L_x^{4+}} \|a_2\|_{X^{s,b}} \|a_3\|_{X^{s,b}} \|h\|_{X^{0,1-b}},
\end{aligned} \tag{4.32}$$

which is summable using Cauchy–Schwarz in $N_2 \sim N_1$.

When $N_2 \ll N_1$, we dyadically decompose h into h_{N_4} and note we must have $N_1 \sim N_4$. We use bilinear Strichartz between F_1 and a_3, and we put $a_2 \in L_{t,x}^{4+}$ to obtain

$$\sum_{N_4 \sim N_1 \gg N_2 \ge N_3} N_1^s \left(\frac{N_3}{N_1}\right)^{1/2} N_3^{-s} \|a_3\|_{X^{s,b}} \|a_2\|_{X^{\epsilon,b}} \|h_4\|_{X^{0,1-b}} \|F_1\|_{L_t^\infty L_x^2},$$

which is again summable. Note that we do lose an extra $\lambda^{-\epsilon}$ in the term $\|a_2\|_{X^{\epsilon,b}}$.

Case 5: $\|a_{N_1} F_{N_2} a_{N_3}\|$. We estimate as in the previous case, but do not need to transfer regularity from the function at the lowest frequency to the highest.

Case 6: $\|a_{N_1} a_{N_2} F_{N_3}\|$. We estimate as in the previous case, but do not need to transfer regularity from the function at the lowest frequency to the highest.

Case 7: $\|a_{N_1} a_{N_2} a_{N_3}\|$. As in standard deterministic local theory.

4.1. Estimate of (4.6)

The estimate in (4.6) essentially follows directly from (4.5), but we sketch the argument. We will again write down the Duhamel formula of (4.9), and apply the I-operator $I_{N(\tau_k^j)}$ on both sides, and estimate

$$\left\| I_N |F + a|^2 (F + a) \right\|_{X^{s,b-1}}. \tag{4.33}$$

We distinguish four different scenarios:

(i) Three random pieces $|F|^2 F$.

(ii) Two random pieces terms, for example $F\bar{F}a$.

(iii) Terms with at least two copies of a, and the highest frequency is on a, for example, the term $a_1 \bar{F}_2 a_3$.

(iv) Terms where F is at the highest frequency $F_1 \bar{a}_2 a_3$.

For situations (i) and (ii), observe that I_N will send $X^{s,b}$ into $X^{1,b}$ by losing

$$N(t_k)^{1-s} \sim \lambda(t_k)^{1-s} \lambda(t_k)^{(1-s)\delta}.$$

Using this estimate directly will miss the desired result by $\lambda(t_k)^{(1-s)\delta}$, and we now detail how to recover this loss.

In case (i), we see from the previous arguments for Terms 1–3 that one beats the desired estimates by $\lambda(t_k)^{-s}$, hence choosing $0 < \delta \ll 1$ small suffices.

In case (ii), estimates of the form (4.26) are applied and one gains a positive power of $\lambda(t)$, for example $\lambda(t_k)^{1/100}$. Such gains are already enough to compensate the $\lambda(t_k)^{(1-s)\delta}$ loss since $0 < \delta \ll 1$.

In case (iii), the estimate follows from standard deterministic arguments, and since the highest frequency is on a, thus $I_N(a_1 \overline{b_2} b_3)$ (where $b = a$ or F) can be estimated (effectively) as $(I_N a_1)\overline{b_2} b_3$, and a standard persistence of regularity argument can close the estimates.

In case (iv), we are only concerned with the situation $N_2 \ll N_1$, and further one only needs to consider $N_1 \geq N = N(\tau_k^j)$. One can distinguish two subcases:

- $N_2 \geq N_1 \lambda(t_k)^{\tilde{\epsilon}}$ (it will be clear soon how we should choose this $\tilde{\epsilon}$),
- $N_2 \leq N_1 \lambda(t_k)^{\tilde{\epsilon}}$.

In the first subcase, we again use persistence of regularity and transfer $\langle D \rangle I_N$ to a_2 by losing $(N_1/N_2)^s$, and one will be able to close (recalling that an error of $\lambda(t_k)^{-\epsilon_1}$ is allowed) if

$$\tilde{\epsilon} s \lesssim \tfrac{1}{10} \epsilon_1. \tag{4.34}$$

Note that the existence $\tilde{\epsilon}$ satisfying (4.35) and (4.34) requires $\epsilon_1 \geq \delta s^2$, which is acceptable. In the second subcase, one follows the same computations as with the term $F_{N_1} a_{N_2} a_{N_3}$ in Case 4, and the bilinear Strichartz estimate gives us an extra $(N_2/N_1)^{1/4}$ $\lesssim \lambda(t_k)^{-\tilde{\epsilon}}$. We will again use the fact I_N will send $X^{s,b}$ into $X^{1,b}$ by losing

$$N(t_k)^{1-s} \sim \lambda(t_k)^{1-s} \lambda(t_k)^{(1-s)\delta},$$

and we are able to close the estimates provided

$$\tilde{\epsilon} \geq 10\delta s. \tag{4.35}$$

This concludes the proof. ■

5. Energy estimates

In this section, we combine the improved probabilistic local well-posedness of Section 4 with the log-log bootstrap scheme, in particular (3.33), to prove the analogue of [20, Proposition 3.1]. We still follow an I-method scheme, but our implementation has two main differences compared to [20]:

- Our LWP theory is different from the standard $H^s(\mathbb{R}^2)$ LWP in [20].
- The function a will play the role of full solution u in [20], and in particular a does not solve the standard NLS, but rather a forced equation with random forcing terms, for which we need to incorporate extra random data type techniques into the I-method computation.

We note that we also take this opportunity to simplify certain aspects of the I-method arguments from [20] in the current setting. Due to the fact that we ultimately combine the energy estimates with the log-log bootstrap, it seems unnecessary to exploit the full cancellation of the I-operator.

Recall that we use the ansatz (3.20). Let $J_{N(t)}$ denote the Fourier multiplier such that

$$J_{N(t)} + I_{N(t)} = \mathrm{Id}. \tag{5.1}$$

Following [20], let

$$\Xi(t) = \frac{\lambda^2}{2} \int |\nabla J_{N(t)} a(0)|^2 \, dx, \tag{5.2}$$

In the rest of the article, we will take $p = \infty-$, and we always assume a small probability set has already been dropped so that (2.5) , (2.12) and (2.13) hold, and for every LWP interval $[\tau_k^j, \tau_k^{j+1}]$, Lemma 4.2 holds. Since we discussed these issues thoroughly in the previous section, we do not revisit them again. We will establish the following result.

Proposition 5.1. *Restricting the subset so that* (2.5), (2.12), (2.13) *and Lemma* 4.2 *hold, we have the following: there exists some* $\alpha_1 > 0$ *such that for all* $t \in [0, T]$*, one has*

$$\left| E(I_{N(t)} a) + \frac{1}{\lambda(t)^2} \Xi(t) \right| \lesssim \left(\frac{1}{\lambda(t)} \right)^{2-\alpha_1}, \tag{5.3}$$

$$|P(I_{N(t)} a(t))| \lesssim \left(\frac{1}{\lambda(t)} \right)^{1-\alpha_1}. \tag{5.4}$$

Remark 5.2. The exact value of α_1 is somewhat different in our setting compared to [20]. Indeed, recall that the authors of [20] establish a result for every $s > 0$, and thus they have a choice of α_1 for each such s. In contrast, we *choose* some $0 < s \ll 1$, and only need to find one such α_1 for this particular s.

Proof of Proposition 5.1. We will focus on estimate (5.3) and, as we will remark, (5.4) follows in a similar (if not simpler) manner. We will only handle the case $t = T$ in Proposition 5.1, and we will denote $N = N(T)$. The proof of (5.3) has two parts:

- an initial estimate

$$\left| E(I_N(a(0))) + \frac{1}{\lambda(T)^2}\Xi(T) \right| \lesssim \left(\frac{1}{\lambda(T)}\right)^{2-\alpha_1} \tag{5.5}$$

  for some $\alpha_1 > 0$, and
- a growth estimate

$$|E(I_N a(T)) - E(I_N a(0))| \lesssim \left(\frac{1}{\lambda(T)}\right)^{2-\alpha_1} \tag{5.6}$$

  for some $\alpha_1 > 0$.

The initial estimate (5.5) follows from the bootstrap assumptions (3.32) and the fact that the potential energy is subcritical compared to the kinetic energy, and one can argue exactly as in the proof of [20, (3.24)]. Thus, the rest of this section is mainly devoted to the proof of (5.6). Recalling (3.26), we compute

$$\begin{aligned}\partial_t E(I_N a) &= \Im \int \overline{I_N \Delta a + |I_N a|^2 I_N a} \left[I_N\big(|a+F|^2(a+F)\big) - |I_N a|^2 I_N a \right] \\ &= A_I + A_{II} + B_I + B_{II} \end{aligned} \tag{5.7}$$

where

$$\begin{aligned} A_I &:= \Im \int \overline{I_N \Delta a} \left[I_N(|a+F|^2(a+F)) - |I_N(a+F)|^2 I_N(a+F) \right], \\ A_{II} &:= \Im \int \overline{I_N \Delta a} \left[|I_N(a+F)|^2 I_N(a+F)) - |I_N a|^2 I_N a \right], \\ B_I &:= \Im \int \overline{|I_N a|^2 I_N a} \left[I_N(|a+F|^2(a+F)) - |I_N(a+F)|^2 I_N(a+F) \right], \\ B_{II} &:= \Im \int \overline{|I_N a|^2 I_N a} \left[|I_N(a+F)|^2 I_N(a+F)) - |I_N a|^2 I_N a \right]. \end{aligned}$$

We will estimate each term separately, up to an observation on cancellation between terms that will be useful in what follows. Indeed, we note that the term $-|I_N F|^2 I_N F$ in A_I will cancel the same term in A_{II}, and the same cancellation also holds between B_I and B_{II}. It is not immediately clear whether such cancellation is crucial, but it simplifies the analysis considerably because subtle probabilistic arguments (as illustrated in the previous section) have to be applied to analyze $|F|^2 F$, resulting in extra subsets of small probability needing to be dropped. In order to redo the same estimates for all N, the analysis is not only more technical, but one needs to be careful about summability of the probabilities of these subsets, and the aforementioned cancellation frees us from this issue.

We will now estimate the terms A_I, A_{II}, B_I, B_{II}.

Estimates of A_I. To estimate

$$A_I := \Im \int \overline{I_N \Delta a} \left[I_N(|a+F|^2(a+F)) - |I_N(a+F)|^2 I_N(a+F) \right], \tag{5.8}$$

we will estimate the integral of A_I within each LWP interval $[\tau_k^j, \tau_k^{j+1}]$ for $k_0 \le k \le k_1$, $1 \le j \le J_k$ and then sum the resulting estimates. Due to the fact that our method is of sub-critical nature, we need to beat the trivial estimate by at least $\lambda(t_k)^{-\delta}$ within $[\tau_k^j, \tau_k^{j+1}]$. With this in mind, we recall the parameters (4.1), and the fact that any loss of $\lambda(t_k)^{-C\epsilon_2}$ or $\lambda(t_k)^{-C\delta s}$ will be acceptable and can be neglected; we do not repeat this point later in the analysis.

We will see from the proof that if one fixes k, the estimate can be performed identically for different j. This follows since there are at most $k\lambda(t_k)^{-\epsilon_2} \lesssim \lambda(t_k)^{-2\epsilon_2}$ LWP intervals. Hence, we can estimate a single LWP interval within $[t_k, t_{k+1}]$ and absorb the loss stemming from counting the number of intervals.

Finally, one will observe that the estimate of $|A_I|$ is monotone in k; it is indeed enough to compute its integral in the last LWP interval $[\tau_{k_1}^{J_k-1}, \tau_{k_1}^{J_k}]$ since $k_1 \sim \log \frac{1}{\lambda(t_k)}$, and a loss of k_1 is also allowed by the previous analysis. At the heuristic level, one may compare it to summing up a geometric series, where the value of the sum is determined by the last term (up to an allowable error).

We apply a Littewood–Paley decomposition to the quadrilinear term (5.8), with frequencies $\xi_i \sim N_i$. We assume that $N_2 \ge N_3 \ge N_4$, and we write

$$a_i = a_{N_i} := P_{N_i} a,$$

and similarly for F_i. We will explicitly estimate two types of terms which are the most difficult ones. Other terms can be estimated via essentially the same (if not easier) analysis. As mentioned above, we will also exploit cancellation which enables us to handle some of the terms with three random pieces. We finally point out that when there are no random terms, one can just follow [20].

One random piece. The most difficult case is when the random term is at the highest allowable frequency, N_2. We consider one local well-posedness interval $I = [\tau_k^j, \tau_k^{j+1}]$, recall $|I| \lesssim \lambda(t_k)^2$, and estimate

$$\sum_{N_1, N_2, N_3, N_4} \int_{\tau_k^j}^{\tau_k^{j+1}} \int \overline{I_N \Delta a_1} \left[I_N(F_2 \overline{a}_3 a_4) - I_N F_2 \overline{I_N a_3} I_N a_4 \right] dx\, dt. \tag{5.9}$$

In order for the integral to be nonzero, we necessarily have $N_2 \gtrsim N$. We handle two cases:

- $N_1 \sim N_2$,
- $N_1 \ll N_2$.

Without loss of generality, assume that $N_i \geq 1$ for all i. When $N_1 \sim N_2$, we will estimate the integral by estimating the term[21]

$$\Bigg|\sum_{N_1\sim N_2\geq N_3\geq N_4} \iint N_1^2\left(\frac{N^{1-s}}{N_1^{1-s}}a_1\right)\left(\frac{N^{1-s}}{N_2^{1-s}}F_2\right)\overline{a_3}a_4\Bigg|. \tag{5.10}$$

We note that the complex conjugate will not be relevant and the quotients come from the definition of the I_N operator. We use Hölder's inequality to get the bound

$$N^{2-2s}\sum_{N_2\sim N_1\geq N_3\geq N_4} N_1^{2s}\|a_1a_3\|_{L^2_{x,t}}\|F_2a_4\|_{L^2_{t,x}}. \tag{5.11}$$

First we sum over a_4 via the triangle inequality, noting that $\|a_4\|_{L_t^\infty L_x^{2+}} \lesssim N_4^{-s+}\lambda(t_k)^{-2s}$. As previously mentioned, we neglect any loss of the form $\lambda(t_k)^{-\epsilon_1}$. Noting that in the current case $N_1 \sim N_2$, and recalling that $F_2 \in L^p_{t,x} = L^{\infty-}_{t,x}$ by (2.13), we find the bound

$$\begin{aligned}
&\lesssim N^{2-2s}\sum_{N_1\geq N_3\geq N_4} N_1^{2s}\|a_1a_3\|_{L^2_{t,x}}\|F_2\|_{L^{\infty-}_{t,x}}|I|^{1/2-}\\
&\sim N^{2-2s}\lambda(t_k)^{1-2s-}\sum_{N_1\geq N_3} N_1^{2s}\|a_1a_3\|_{L^2_{t,x}}\|F_2\|_{L^p_{t,x}}.
\end{aligned} \tag{5.12}$$

Finally, we use the bilinear Strichartz esimates to obtain the estimate

$$\lesssim N^{2-2s}\lambda(t_k)^{1-}\sum_{N_1\geq N_3} N_1^s\|a_1\|_{X^{0,b}}\|F_2\|_{L^p_{t,x}}N_3^s\|a_3\|_{X^{0,b}}\left(\frac{N_3}{N_1}\right)^{1/2-s}. \tag{5.13}$$

Since both $N_1^s\|a_1\|_{X_{0,b}}$ (by definition) and $\|F_2\|_{L^p_{t,x}}$ (by (2.13)) are ℓ_2 summable, we may apply the Cauchy–Schwarz inequality between these terms, and the triangle inequality to sum over $N_3 \leq N_1$. We can ultimately estimate the contribution of this term to (5.13) by

$$N^{2-2s}\lambda(t_k)^{1-2s-}.$$

Summing over all LWP intervals and applying (3.33), (4.2) (recall also (4.1)), one has

$$\left|\int_0^T A_I\right| \lesssim \sum_{k_0}^{k_1} kN^{2-2s}\lambda(t_k)^{1-2s-}\lambda(t)^{-\epsilon_2} \lesssim N^{2-2s} \lesssim \lambda(T)^{-2(1+\delta)(1-s)} \lesssim \lambda(T)^{-2+s}, \tag{5.14}$$

which is the desired estimate.

We also record the following simple observation from the computation above as a remark to reference later in the proof. We will not repeat the same argument later.

[21] Strictly speaking, (5.9) is not bounded by (5.10), but by (5.11). What we mean here is that one can think about the estimate of (5.9) as the estimate of (5.10), thus naturally leading to the estimate of (5.11).

Remark 5.3. Provided we can estimate

$$\int_{\tau_k^j}^{\tau_k^{j+1}} A_I \lesssim N^{1-c_0 s}$$

for some $c_0 > 0$, for example $c_0 = \frac{1}{100}$, we are able to sum the estimates up along all the LWP intervals.

Next we turn to the case when $N_1 \ll N_2$; we first observe that necessarily $N_3 \sim N_2$. For notational convenience, we will use

$$S := X^{0,b}$$

in the rest of the section. We discuss two subcases:

- $N_1 \geq N$,
- $N_1 \leq N$.

First if $N_1 \geq N$, we may reduce to estimating

$$\begin{aligned}
&\Big| \sum_{N_3 \sim N_2 \geq N_1, N_4} \iint N_1^2 \Big(\frac{N^{1-s}}{N_1^{1-s}} a_1 \Big) \Big(\frac{N^{1-s}}{N_1^{1-s}} F_2 \Big) a_3 a_4 \Big| \\
&\quad \lesssim N^{2-2s} \Big| \sum_{N_3 \sim N_2 \geq N_1, N_4} \iint N_1^{2s} a_1 F_2 a_3 a_4 \Big| \\
&\quad \lesssim N^{2-2s} \sum_{N_3 \sim N_2 \geq N_1, N_4} \lambda(t_k)^{1-\varepsilon} N_1^{2s} \|a_1\|_S \|a_3\|_S N_2^{\varepsilon} \|F_2\|_{L^{\infty-}_{t,x}} N_4^{-s} \|a_4\|_S \\
&\quad \lesssim N^{2-2s} \lambda(t_k)^{1-\varepsilon} \sum_{N_3 \sim N_2 \geq N_1} N_1^{2s} \|a_1\|_S N_2^{\varepsilon} \|a_3\|_S. \qquad (5.15)
\end{aligned}$$

Note that in the first line of (5.15), we either estimate $I_N(F_2 a_3 a_4)$ whose output frequency lies in $|\xi| \sim N_1$, or we estimate $I_N F_2 I_N a_3 I_N a_4$, which has an I-operator smoothing at frequency $\sim N_2 \geq N_1$. Ultimately we again obtain the bound

$$N^{2-2s} \lambda(t_k)^{1-\varepsilon-2s},$$

which is enough from Remark 5.3.

Finally, when $N_1 \leq N$, then we may reduce to estimating

$$\Big| \sum_{N_3 \sim N_2 \geq N_1, N_4,\, N \geq N_1} \iint N_1^2 a_1 F_2 N^{1-s} N_2^{-1+s} a_3 a_4 \Big|, \qquad (5.16)$$

and in this case, one ends up with $\lambda(t_k)^{1-2s-} N^{2-s}$, which is sufficient by Remark 5.3.

Two random pieces. We estimate the term with a_1, F_2, F_3, a_4. Recall again that unless $N_2 \gtrsim N$ the expression we are estimating is zero. Once again, we handle two cases:

- $N_1 \sim N_2$,
- $N_1 \ll N_2$.

When $N_1 \sim N_2$, we estimate the integral via

$$\sum_{N_2 \sim N_1 \geq N_3 \geq N_4} \int N_1^2 \left(\frac{N^{1-s}}{N_1^{1-s}} a_1\right)\left(\frac{N^{1-s}}{N_2^{1-s}} F_2\right) F_3 a_4$$
$$= N^{2-2s} \sum_{N_2 \sim N_1 \geq N_3 \geq N_4} \int N_1^{2s} a_1 F_2 F_3 a_4$$
$$\lesssim N^{2-2s} \sum_{N_2 \sim N_1 \geq N_3 \geq N_4} N_1^{2s} \|a_1 a_4\|_{L^2_{t,x}} \|F_2 F_3\|_{L^2_{x,t}}. \tag{5.17}$$

We apply bilinear Strichartz and use the $L^{\infty-}_{t.x}$ control from F_{N_1} (using randomness) to have the above controlled by

$$\lesssim N^{2-2s} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1^s \|a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|F_3\|_{L_t^{\infty-} L_x^4} |I|^{1/2-} N_4^s \left(\frac{N_4}{N_1}\right)^{1/2-s} \|a_4\|_S$$
$$\lesssim N^{2-2s} \lambda(t_k)^{1-} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1^s \|a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|F_3\|_{L_t^{\infty-} L_x^4} \left(\frac{N_4}{N_3}\right)^{1/2-s} N_4^s \|a_4\|_S. \tag{5.18}$$

Applying Cauchy–Schwarz in N_4, one derives the estimate

$$\lesssim N^{2-2s} \lambda(t_k)^{1-} \sum_{N_1 \sim N_2 \geq N_3} N_1^s \|a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|F_3\|_{L_t^{\infty-} L_x^4} \|a_3\|_S \left(\frac{N_3}{N_1}\right)^{1/2-s}$$
$$\lesssim N^{2-2s} \lambda(t_k)^{1-} \sum_{N_1 \sim N_2 \geq N_3} N_1^s \|a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|F_3\|_{L_t^{\infty-} L_x^4} \left(\frac{N_3}{N_1}\right)^{1/2-s}. \tag{5.19}$$

To conclude, we sum over N_3, and then apply Cauchy–Schwarz in $N_1 \sim N_2$, which yields the bound

$$N^{2-2s} \lambda(t_k)^{1-}.$$

Next, to estimate the expression when $N_1 \ll N_2$, we again note that necessarily $N_2 \sim N_3$. As above, we split into subcases:

- $N_1 \geq N$,
- $N_1 \leq N$.

In the first subcase, one estimates

$$N^{2-2s} \sum_{N_3 \sim N_2 \geq N_1, N_4} \int N_1^{2s} a_1 F_2 F_3 a_4$$
$$\lesssim N^{2-2s} \sum_{N_3 \sim N_2 \geq N_1, N_4} \|F_2 F_3\|_{L^2_{x,t}} N_1^s \|a_1\|_S N_4^s \|a_4\|_S \min\left(\left(\frac{N_1}{N_4}\right)^s, \left(\frac{N_4}{N_1}\right)^{1/2-s}\right). \tag{5.20}$$

As in the first case of A_I, we handle the double sum over N_1, N_4, which we may estimate by

$$N^{2-2s} \sum_{N_3 \sim N_2} \|F_2 F_3\|_{L^2_{t,x}} \lesssim N^{2-2s} \sum_{N_2} |I|^{1/2-} \|F_2\|_{L_t^{\infty-} L_x^4} \|F_3\|_{L_t^{\infty-} L_x^4}$$
$$\lesssim N^{2-2s} \lambda(t_k)^{1-}. \tag{5.21}$$

Finally, the case $N_1 \leq N$ proceeds analogously with an extra N^s loss, which is allowable.

Three random pieces. As mentioned in our discussion of the cancellation of three random terms above, we only need to control

$$\int_I \int \overline{\Delta I_N a} I_N(|F|^2 F). \tag{5.22}$$

Here, we recall (4.13) and bound the above by

$$\|I_N a\|_{X^{1,b}[I]} \|I_N(|F|^2 F)\|_{X^{1,1-b}[I]} \lesssim \|I_N a\|_{X^{1,b}[I]} N^{1-s}. \tag{5.23}$$

Note we use the fact that I_N can gain $1-s$ derivatives by losing N^{1-s}. Now, plugging in (4.7), we bound the above by

$$N(T)^{1-s} \left(\frac{N(T)}{N(t_k)} \right)^{1-s} \left(\frac{1}{\lambda(t_k)} \right)^{1+\epsilon_1} \tag{5.24}$$

Summing over all LWP intervals yields the bound

$$N(T)^{1-s} \frac{1}{\lambda(t)^{1+C\epsilon_2}},$$

which is acceptable provided $0 < \delta \ll s$, and $0 < \epsilon_2 \ll 1$.

Estimates of A_{II}. We recall the expression for A_{II}:

$$A_{II} := \Im \int \overline{I_N \Delta a} \left[|I_N(a+F)|^2 I_N(a+F)) - |I_N a|^2 I_N a \right].$$

Also recalling again our discussion on the cancellation of the three random terms, we note there will be no need to consider the case of three random pieces here. In light of Remark 5.3, we will work on $I = [\tau_k^j, \tau_k^{j+1}]$, and prove an estimate on this interval.

One or two random pieces. We may now combine the estimates for one or two random pieces. As above, we let $N_2 \geq N_3 \geq N_4$. Once again, we treat the case where the random piece is at the highest allowable frequency. We consider the cases:

(1) $N_1 \sim N_2 \geq N_3 \geq N_4$ and
- $N_1 \geq N$,
- $N_1 \leq N$.

(2) $N_1 \ll N_2$ (in this case one must have $N_2 \sim N_3$) and
- $N_1 \geq N$ (in this case one must have $N_2 \geq N$),
- $N_1 \leq N$.

Recall that

$$\|\nabla I a\|_S \lesssim \left(\frac{N}{N(t_k)}\right)^{1-s} \frac{1}{\lambda(t_k)} \tag{5.25}$$

and

$$\|a\|_{X^{s,b}} \sim \|\nabla^s a\|_S \sim \frac{1}{\lambda(t_k)^s}. \tag{5.26}$$

We start with the subcase $N_1 \sim N_2 \geq N_3 \geq N_4$, $N_1 \geq N$. In this case, we estimate

$$\begin{aligned} \sum_{N_2 \sim N_1 \geq N_3 \geq N_4} \int N_1^2 I a_1 \frac{N^{1-s}}{N_1^{1-s}} F_2 I a_3 I a_4 \\ \lesssim N^{1-s} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1 N_1^s \|I a_1 I a_4\|_{L^2_{x,t}} \|F_2 I a_3\|_{L^2_{t,x}}. \end{aligned} \tag{5.27}$$

Using bilinear Strichartz estimates for the a_1, a_4 term, and the random data control for F_2, we derive

$$\begin{aligned} &\lesssim N^{1-s} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1 N_1^s (N_4/N_1)^{1/2} \|I a_1\|_S \|a_4\|_S \|F_2\|_{L^4_{t,x}} \|a_3\|_S \\ &= N^{1-s} \lambda(t_k)^{\frac{1}{2}-} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1 \|I a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|a_3\|_S (N_4/N_1)^{1/2-s} N_4^s \|a_{N_4}\|_S \\ &\lesssim N^{1-s} \lambda(t_k)^{\frac{1}{2}-} \sum_{N_1 \geq N_3} N_1 \|I a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|a_3\|_S \|a_3\|_{X^{s,b}} (N_3/N_1)^{1/2-s} \\ &\lesssim N^{1-s} \lambda(t_k)^{\frac{1}{2}-} \frac{1}{\lambda(t_k)^s} \sum_{N_1 \geq N_3} N_1 \|I a_1\|_S \|F_2\|_{L_t^{\infty-} L_x^4} \|a_3\|_S (N_3/N_1)^{1/2-s} \\ &\lesssim N^{1-s} \lambda(t_k)^{\frac{1}{2}-} \frac{1}{\lambda(t_k)^s} \|\nabla I a\|_{S_0} \\ &\lesssim N^{1-s} \lambda(t_k)^{\frac{1}{2}-} \frac{1}{\lambda(t_k)^s} \frac{N^{1-s}}{N(t_k)^{1-s}} \frac{1}{\lambda(t_k)}. \end{aligned} \tag{5.28}$$

This is desirable. It should be remarked that we do not use any regularity of a_{N_3}, so the above arguments also work when a_{N_3} is replaced by F_{N_3}.

Now, we turn to the subcase $N_1 \sim N_2 \geq N_3 \geq N_4$, $N_1 \leq N$; then the I-operator is just the identity map. We estimate

$$\begin{aligned} \sum_{N \geq N_1 \sim N_2 \geq N_3 \geq N_4} \int N_1^2 a_1 F_2 a_3 a_4 \\ \lesssim \lambda(t_k)^{\frac{1}{2}-} \sum_{N_1 \sim N_2 \geq N_3 \geq N_4} N_1^2 \|a_1\|_S \|F_2\|_{L_t^{-\infty} L_x^4} \|a_3\|_S \|a_4\|_S (N_4/N_1)^{1/2} \\ \lesssim N^{2-s} \lambda(t_k) \frac{1}{\lambda(t_k)^s}. \end{aligned} \tag{5.29}$$

We note that while we need s derivatives of a_1, no regularity of a_3 is used and thus this argument applies equally to the case with two random pieces.

Now, let us turn to the subcase $N_1 \ll N_2$; then one must have $N_2 \sim N_3$. We first consider the subcase $N_1 \geq N$ then necessarily $N_2 \geq N$. We observe that

$$\|I_N F_2\|_{L^\infty_{t,x}} \lesssim N_2^{s-1} N^{1-s} \|F_2\|_{L^\infty_{t,x}}. \tag{5.30}$$

One may estimate

$$\sum_{N_2 \sim N_3 \geq N \geq N_1, N_4} N^{2-2s} \frac{N_1^{2-2s}}{N_2^{2-2s}} \int N_1^{2s} a_1 F_2 a_3 a_4$$
$$\lesssim \lambda(t_k) \|F_2\|_{L^\infty_{t,x}} \|a_3\|_S (N_1/N_2)^{2-2s} N_1^{2s} \|a_1 a_4\|_{L^2_{t,x}}, \tag{5.31}$$

and we further estimate

$$N_1^{2s} \|a_1 a_4\|_{L^2_{t,x}} \lesssim N_1^s N_4^s \|a_1\|_S \|a_4\|_S \min\big((N_4/N_1)^{1/2-s}, (N_1/N_4)^s\big). \tag{5.32}$$

Plugging this back to (5.31), one finishes with the bound $\lambda(t_k)^{1-2s} N^{2-2s}$.

Finally, we are left with the case $N_1 \ll N_2$, $N_3 \sim N_2$, $N_1 \leq N$; one simply estimates this expression as

$$\sum_{N_1, N_2 \sim N_3, N_4} N_1^2 a_1 F_2 a_3 a_4 \lesssim N^{2-s} \lambda(t_k)^{1-s}, \tag{5.33}$$

which is sufficient.

The estimate with two random terms follows as for A_I.

Estimates for B_I and B_{II}. These estimations proceed similarly to the previous ones, but are somewhat simpler since we do not lose derivatives, and indeed it is easy to see that when there are at least three random pieces in the estimate, the proof becomes more or less trivial. This is in sharp contrast to the case of A_I and A_{II}. Also, the purely deterministic case follows from the estimates in [20].

We recall

$$B_I := \Im \int \overline{I_N(|a|^2 a)} \big[I_N(|a+F|^2(a+F)) - |I_N(a+F)|^2 I_N(a+F)\big],$$
$$B_{II} := \Im \int \overline{I_N(|a|^2 a)} \big[|I_N(a+F)|^2 I_N(a+F)) - |I_N a|^2 I_N a\big].$$

One random piece. We will record the estimates involving one random piece. Once again we let $N_1 \geq N_2 \geq N_3$ and $N_4 \geq N_5 \geq N_6$. We will see that B_I and B_{II} are handled in a similar manner to A_I and A_{II} and hence we will sketch the estimates for B_I, and leave B_{II} to the interested reader. We ignore complex conjugates as they will not feature in our argument.

We consider

$$\int I_N a_1 I_N a_2 I_N a_3 \left[I_N(F_4 a_5 a_6) - I_N F_4 I_N a_5 I_N a_6\right], \tag{5.34}$$

and without loss of generality, assume that $N_i \geq 1$ for all i. As previously, we note that in order for this expression to be nonzero, we will need $N_4 \gtrsim N$. We let N_{123} be the resulting frequency from the convolution of the first three terms. In this setting, we need to consider two cases:

- $N_1 \gtrsim N_4$,
- $N_1 \ll N_4$.

In the first case, we use Bernstein's and Hölder's inequalities to estimate

$$\begin{aligned}
&\Big|\sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} \iint \Big(\frac{N^{1-s}}{N_1^{1-s}} a_1\Big) a_2 a_3 \Big(\frac{N^{1-s}}{N_4^{1-s}} F_4\Big) a_5 a_6\Big| \\
&\simeq N^{2-2s} \sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} N_1^{s-1} N_4^{s-1} \|a_1 a_2 a_3 a_5\|_{L^2_{t,x}} \|F_4 a_6\|_{L^2_{t,x}} \\
&\simeq N^{2-2s} \sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} N_1^{2s-2+1/2} \|a_1 a_2 a_3 a_5\|_{L^2_t L^1_x} \|F_4 a_6\|_{L^2_{t,x}} \\
&\simeq N^{2-2s} \sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} N_1^{2s-2+3/2} \|a_2\|_{L^\infty_t L^2_x} \|a_3\|_{L^\infty_t L^2_x} \|a_1 a_5\|_{L^2_{t,x}} \|F_4 a_6\|_{L^2_{t,x}}.
\end{aligned} \tag{5.35}$$

Since $2s - 1/2 < 2s$, we can estimate this as in the kinetic term.

In the second case, if $N_1 \ll N_4$, then since the convolution of the first three terms and the convolution of the last three terms are paired, we must have $N_4 \sim N_5$, and we can further estimate based on whether

- $N_1 \geq N$,
- $N_1 \leq N$,

as before.

Two random pieces. There are two subcases to consider:

- $N_1 \gtrsim N_4$,
- $N_1 \ll N_4$.

We again estimate mimicking the kinetic term estimates, to obtain

$$\begin{aligned}
&\Big|\sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} \iint \Big(\frac{N^{1-s}}{N_1^{1-s}} a_1\Big) a_2 a_3 \Big(\frac{N^{1-s}}{N_4^{1-s}} F_4\Big) F_5 a_6\Big| \\
&\qquad \simeq N^{2-2s} \sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} N_1^{s-1} N_4^{s-1} \|a_1 a_2 a_3 a_6\|_{L^2_{t,x}} \|F_4 F_5\|_{L^2_{t,x}} \\
&\qquad \simeq N^{2-2s} \sum_{N_4 \sim N_1 \geq N_2, N_3, N_5, N_6} N_1^{2s-2+1/2} \|a_1 a_2 a_3 a_6\|_{L^2_t L^1_x} \|F_4 F_5\|_{L^2_{t,x}}
\end{aligned} \tag{5.36}$$

and again we can use Bernstein on a_2, a_3.

In the second case, if $N_1 \ll N_4$, then since the convolution of the first three terms and the convolution of the last three terms are paired, we must have $N_4 \sim N_5$, and we can further estimate based on whether

- $N_1 \geq N$,
- $N_1 \leq N$,

and again we can argue as for the kinetic terms.

Estimate for momentum. These estimates proceed via direct computation, and we refer as well to the explanation in [20]; the proof of (5.4) is similar to the proof for the kinetic part of (5.4). ■

6. Proof of the bootstrap lemma and the main theorem

In this section we establish the main bootstrap result, Lemma 3.6, as well as the main Theorem 3.1. First, we recall our ansatz (3.20):

$$u(t,x) = a(t,x) + F(t,x), \tag{6.1}$$

$$a(t,x) = \frac{1}{\lambda(t)}(Q_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{-i\gamma(t)}, \tag{6.2}$$

where a plays the role of the full solution u in [20], and satisfies the forced NLS (3.26), and where the parameters $\lambda(t), x(t), b(t), \gamma(t)$ are chosen so that the orthogonality conditions (3.21)–(3.24) hold. Once we establish the desired energy estimates for $E(I_N a)$ and $P(I_N a)$ under the bootstrap assumptions of Lemma 3.6, the proof of Lemma 3.6 essentially follows as in [20, Section 4], with some changes in our current setting which we highlight below. In particular, we will verify that given our estimates on $E(I_N a)$ and $P(I_N a)$, the key computations in [36–38, 40] still hold following the bootstrap scheme in [45].

It should be noted that unlike the full solution u, the nonlinear component of the solution a does not satisfy an exact mass conservation law, which adds additional technical difficulties in the last step of Section 6.1 below.

6.1. *Energy estimates imply persistence of log-log regime*

Overview. There are five main steps in this subsection. In Step 1, one works with the rescaled time variable to obtain the modulation equations for ϵ. In Step 2, one obtains some control of ϵ by modulation analysis and analysis of the almost conserved quantities $P(I_N a)$ and $E(I_N a)$. Steps 1 and 2 form the basis of log-log analysis. We subsequently analyze the modulation equations for ϵ and corresponding estimates obtained in Step 2. In the crucial Step 3, one recovers the key local virial estimate of Merle and Raphaël, which will be responsible for the upper bound of the log-log blowup rate. In Steps 4 and 5, one controls the L^2 dispersion at infinity and explores the fact that a almost enjoys a mass conservation law. Finally, one recovers the Lyapunov type control of Merle and Raphaël, which is responsible for the lower bound of the log-log blowup rate.

Step 1. We use the rescaled time variable s, where $ds = \lambda^{-2} dt$, and we set $t(s_0) = 0$ and $t(s_+) = T$. We use the forced NLS (3.26) to derive

$$\partial_s \Sigma_b + \partial_s \epsilon_1 - M_-(\epsilon) + b\Lambda\epsilon_1 = \left(\frac{\lambda_s}{\lambda} + b\right)\Lambda\Sigma_b + \tilde{\gamma}_s \Theta_b + \frac{x_s}{\lambda}\nabla\Sigma_b + \left(\frac{\lambda_s}{\lambda} + b\right)\Lambda\epsilon_1 + \tilde{\gamma}_s \epsilon_2 + \frac{x_s}{\lambda}\nabla\epsilon_1 + \Im\Psi_b - R_2(\epsilon) - G_2, \tag{6.3}$$

$$\partial_s \Theta_b + \partial_s \epsilon_2 + M_+ + b\Lambda\epsilon_2 = \left(\frac{\lambda_s}{\lambda} + b\right)\Lambda\Theta_b - \tilde{\gamma}_s \Sigma_b + \frac{x_s}{\lambda}\nabla\Theta_b + \left(\frac{\lambda_s}{\lambda} + b\right)\Lambda\epsilon_2 - \tilde{\gamma}_s \epsilon_1 + \frac{x_s}{\lambda}\nabla\epsilon_2 - \Re\Psi_b + R_1(\epsilon) + G_1, \tag{6.4}$$

where $\tilde{\gamma} = -s - \gamma$, and M_+, M_-, and R_1, R_2 are defined via

$$|\tilde{Q}_b + \epsilon|^2(\tilde{Q}_b + \epsilon) - |\tilde{Q}_b|^2 Q_b = M_+(\epsilon) + iM_-(\epsilon) + R_1(\epsilon) + iR_2(\epsilon), \tag{6.5}$$

i.e. $M_\pm$ picks up the first order term (with respect to ϵ), and $R_1 + iR_2$ picks up the second and higher order term (in ϵ).

And $G = G_1 + iG_2$ is defined via

$$G(t,x) = -\big(|\tilde{Q}_b + \epsilon + \tilde{F}|^2(\tilde{Q}_b + \epsilon + \tilde{F}) - |\tilde{Q}_b + \epsilon|^2(\tilde{Q}_b + \epsilon)\big), \tag{6.6}$$

where $\tilde{F}(t,x) = \lambda(t) F(t, \lambda(t)x + x(t)) e^{i\gamma(t)}$.

Note that (6.3) and (6.4) are exactly equations (4.2) and (4.3) in [20] except that we have two extra terms, G_1 and G_2. We will see that these terms can be treated perturbatively due to the fact that F is the linear evolution of randomized initial data.

Step 2. We now derive some preliminary estimates using (almost) conservation laws, and modulation estimates. In particular, using our control of $E(I_N a)$ and $P(I_N a)$ obtained in the previous section, we derive the following result.

Lemma 6.1. *For all $s \in [s_0, s_+]$,*

$$\Big| 2(\epsilon_1, \Sigma_b + b\Lambda\Theta_b - \Re\Psi_b) + 2(\epsilon_2, \Theta_b - b\Lambda\Sigma_b - \Im\Psi_b) - 2\left(2\Xi + \int |I_{N\lambda}\nabla\epsilon|^2 - 3Q^2 I_{N\lambda}\epsilon_1^2 - \int Q^2 I^2_{N\lambda\epsilon_2}\right)\Big| \leq \delta_0\left(\int |\nabla I_{N\lambda}\epsilon|^2 + \int \epsilon^2 e^{-|y|}\right) + \Gamma_b^{1-C\eta}, \tag{6.7}$$

$$|(\epsilon_2, \nabla Q)| \leq \delta_0\left(\int |\nabla I_{N\lambda}\epsilon|^2\right)^{1/2} + \Gamma_b^{10}. \tag{6.8}$$

Here $\delta_0 > 0$ is some small constant. This step is exactly the same as the derivation of (4.5) and (4.6) in [20]. In this step, we rely on the bootstrap assumption (3.31), and the almost conservation law of Proposition 5.1.

By substituting (3.21)–(3.24) into (6.3) and (6.4), we derive the following standard modulation estimate.

Lemma 6.2. *For $s \in [s_0, s_+]$,*

$$\left|\frac{\lambda_s}{\lambda} + b\right| + |b_s| + |x_s| \lesssim \Xi(s) + \int |\nabla I_{N\lambda}\epsilon|^2 + \int \epsilon^2 e^{-|y|} + \Gamma_b^{1-C\eta} + \mathcal{F}(s), \quad (6.9)$$

$$\left|\tilde{\gamma}_s - \frac{(\epsilon_1, L_+\Lambda^2 Q)}{\|\Lambda Q\|_{L_x^2}^2}\right| \lesssim \Gamma_b^{1-C\eta} + \mathcal{F}(s) \quad (6.10)$$

where $\mathcal{F}(s) \geq 0$ satisfies

$$\int_s^{s+} \mathcal{F}(s)\, ds \lesssim \lambda(s)^{\alpha_2}, \quad \forall s \in [s_0, s_+], \quad (6.11)$$

with some $\alpha_2 > 0$.

Remark 6.3. The extra term $\mathcal{F}$ is completely perturbative, though it is only estimated in the time average sense; but as this term appears when estimating the time derivative of the modulation parameters, this is sufficient. Heuristically, pointwise,

$$\mathcal{F}(s) \sim -\partial_s \lambda(s)^{\alpha_2} \sim -\frac{\lambda_s}{\lambda}\lambda^{\alpha_2} \sim b\lambda^{\alpha_2} \ll \Gamma_b^{100}.$$

Lemma 6.2 should be compared with [20, (4.7) and (4.8)] (in the H^s setting). Compared with the standard modulation estimates in the H^1 setting, the term $\mathcal{F}(s)$ is introduced to account for the cut-off $I_{N\lambda}$ in the estimate.

In our setting, we need to verify that the extra term in (6.3) and (6.4) is also perturbative. To see this, we briefly recall how the modulation estimate is done. To derive (6.9) and (6.10), one substitutes the four orthogonality conditions (3.21)–(3.24) into (6.3) and (6.4) to cancel the $\partial_s\epsilon_1, \partial_s\epsilon_2$ terms. For example, to substitute (3.21) into (6.3) and (6.4), one needs to take the L_x^2 inner product of (6.3) and $y^2\widetilde{Q}_b$, and the L_x^2 inner product of (6.4) and $y^2\theta_b$, respectively, and sum up. Compared with [20], we obtain extra terms resulting from G_1, G_2, which satisfy

$$(|y|^2|\widetilde{Q}_b|, |G|) \lesssim \int |y|^2|\widetilde{Q}_b|(|\widetilde{Q}_b + \epsilon|^2|\tilde{F}| + |\widetilde{Q}_b + \epsilon|\,|\tilde{F}|^2). \quad (6.12)$$

We then claim that for any $s_1 \in [s_0, s_+)$, we have

$$\int_{s_1}^{s_+} \int |y|^2|\widetilde{Q}_b|(|\widetilde{Q}_b + \epsilon|^2|\tilde{F}| + |\widetilde{Q}_b + \epsilon|\,|\tilde{F}|^2) \lesssim \lambda(s_1)^{\alpha_2} \quad \text{for some } \alpha_2 > 0, \quad (6.13)$$

and thus these extra terms may be absorbed into $\mathcal{F}$ which satisfies (6.11). To establish this bound, we proceed as follows. Let $t(s_1) \in [t_{k_1}, t_{k_1+1}]$ and $T_+ = s_+$, and $\lambda(T_+) \sim 2^{-k_+}$.

We can split $[t_{k_1}, T_+)$ into disjoint intervals $\{I_k\}_{k=k_1}^{k_+}$, and we may split every I_k into disjoint LWP intervals $I_k^j = [\tau_k^j, \tau_k^{j+1}]$ such that $|I_k^j| \sim \lambda(t_k)^{-2}$. Recall that for any k, there exist at most k such intervals, via the bootstrap assumption (3.33). Now, we may estimate the LHS of (6.13), in the original nonrescaled variable, as

$$
\begin{aligned}
&\int_{s_1}^{s_+} \int |y|^2 |\widetilde{Q}_b| (|\widetilde{Q}_b + \epsilon|^2 \tilde{F} + |\widetilde{Q}_b + \epsilon| \, |\tilde{F}|^2) \\
&\lesssim \int_{t_{k_1}}^{T_+} \frac{1}{\lambda(t)^{1/2}} (\|a\|_{L_x^4}^2 \|F\|_{L_x^4} + \|a\|_{L_x^4} \|F\|_{L_x^4}^2 + \|F\|_{L_{t,x}^4}^3) \, dt \\
&\lesssim \sum_{k=k_1}^{k_{+1}} \sum_j \big(\|a\|_{L_{t,x}^4[I_k^j]}^2 \|F\|_{L_{t,x}^4 I_k^j} + \|a\|_{L_{t,x}^4[I_k^j]} \|F\|_{L_{t,x}^4[I_k^j]}^2 + \|F\|_{L_{t,x}^4[I_k^j]}^3 \big).
\end{aligned} \tag{6.14}
$$

Note that up to an exceptional set of small probability (depending on p), one has

$$
\|F(t,x)\|_{L_{t,x}^p} \lesssim 1, \tag{6.15}
$$

which, combined with the estimate $\|F(t,x)\|_{L_t^\infty L_x^2} \lesssim 1$, gives

$$
\|F(t,x)\|_{L_t^4 L_x^4 [I_k^j]} \lesssim |I_k^j|^{\alpha_p}, \quad \text{where} \quad \lim_{p \to \infty} \alpha_p = 1/4, \tag{6.16}
$$

By the standard local theory,[22]

$$
\|a\|_{L_{t,x}^4[I_k^j]} \lesssim 1, \tag{6.17}
$$

hence we can choose p large enough, and estimate (6.14) by

$$
\sum_{k=k_1}^{k_+} k 2^{-2k\alpha_p} \lesssim \lambda(s_1)^{(-2\alpha_p)+}, \tag{6.18}
$$

which establishes (6.13), and consequently Lemma 6.2. We will repeatedly rely on the above argument to handle the extra terms caused by G_1, G_2; we do not repeat the details.

Step 3. This step mainly concerns the derivation of the (local) virial estimate, as well as its sharpening via the tail term $\tilde{\zeta}_b$. This is the core part of the Merle–Raphaël log-log analysis [36–38, 40]. The key point here, similar to [20], is to make sure the original Merle–Raphaël computation remains valid by showing all extra terms introduced are perturbative.

One has the following virial estimates.[23]

[22] Here we can simply apply the usual deterministic L_x^2 local theory rather than the modified probabilistic version of the current article.

[23] Estimate (1.14) was called the local virial estimate in [36], and the global virial estimate in [20].

Lemma 6.4. *There exists $c_0 > 0$ such that for all $s \in [s_0, s_+)$,*

$$b_s \geq c_0\left(\Xi(s) + \int |\nabla I_{N\lambda}\epsilon|^2 + \int \epsilon^2 e^{-|y|} \right) - \Gamma_b^{1-C\eta} - \mathcal{F}(s), \tag{6.19}$$

where $\mathcal{F}$ satisfies (6.11).

Lemma 6.5. *Let*

$$f_1 := \frac{b}{4}\|y\tilde{Q}_b\|_2^2 + \frac{1}{2}\Im \int y\nabla\tilde{\zeta}_b\overline{\tilde{\zeta}_b} + (\epsilon_2, \Lambda\Re\tilde{\zeta}_b) - (\epsilon_1, \Lambda\Im\tilde{\zeta}_b). \tag{6.20}$$

Then for a universal constant c_1 and all $s \in [s_0, s_+)$,

$$\partial_s f_1(s) \geq c_1\left(\Xi(s) + \int |\nabla I_{N\lambda}\epsilon|^2 + \int \epsilon^2 e^{-|y|} + \Gamma_b \right) - \frac{1}{\delta_1}\int_{A\leq|x|\leq 2A} |\xi|^2 - \mathcal{F}(s), \tag{6.21}$$

where $\mathcal{F}$ satisfies (6.11).

Remark 6.6. In the previous lemma, one should think of f_1 as a modified version of b, in particular satisfying $f_1 \sim \frac{1}{4}\|yQ\|_2^2 b$.

Lemmas 6.4 and 6.5 should be compared with Lemmas 4.3 and 4.4, respectively, in [20]. We can again use the argument from Lemma 6.2 above to argue that the extra terms created by G_1 and G_2 in (6.3) and (6.4) can also be absorbed into the error $\mathcal{F}$. For example, to derive Lemma 6.4, one computes the L_x^2 inner product of $-\Lambda\Theta_b$ and (6.3), and the L_x^2 inner product of $\Lambda\Sigma_b$ and (6.4), and sums them together, substituting into (3.23). Ultimately, the extra terms caused by G_1, G_2 are controlled by

$$(\Lambda\tilde{Q}_b|, |G|) \lesssim \int |\Lambda\tilde{Q}_b|(|\tilde{Q}_b + \epsilon|^2|\tilde{F}| + |\tilde{Q}_b + \epsilon|\,|\tilde{F}|^2), \tag{6.22}$$

which can be handled similarly to (6.12) above. We omit the details.

Step 4. In this step, we need to control the $L_x^2(\mathbb{R}^2)$ dispersion at infinity. Recall that

$$A = A_b = e^{a\pi/b},$$

and Ψ is a radial cut-off function, with $\Psi = 0$ for $|x| \leq 1/2$ and $\Psi = 1$ for $|x| \geq 3$. Let $\Psi_A(x) = \Psi(x/A)$.

Lemma 6.7. *For some universal constants $C, c_3 > 0$ and all $s \in [s_0, s_+)$,*

$$\partial_s \int \Psi_A|\epsilon|^2 \geq c_3 b \int_{A\leq|x|\leq 2A} |\epsilon|^2 - \Gamma_b^{a/2}\int |\nabla I_{N\lambda}\epsilon|^2 - \Gamma_b^{1+Ca} - \mathcal{F}(s) - \partial_s\mathcal{H}(s), \tag{6.23}$$

where $\mathcal{F}$ satisfies (6.11), *and $\mathcal{H}$ satisfies the estimate*

$$|\mathcal{H}(s)| \lesssim \lambda(s)^{\alpha_3} \quad \text{for some } \alpha_3 > 0. \tag{6.24}$$

Lemma 6.7 corresponds to [20, Lemma 4.5], up to certain technical modifications. We quickly go over its proof, focusing only on the differences from the proof in [20]. While the $\mathcal{F}$ term could actually be absorbed into $\partial_s \mathcal{H}$, we choose to proceed in a manner that more closely follows the original presentation of [20].

Remark 6.8. We recall the tail $\tilde{\zeta}_b$ introduced in the previous section, and set $\tilde{\epsilon} = \epsilon - \tilde{\zeta}_b$. Note that $\|\nabla\epsilon - \nabla\tilde{\epsilon}\|_{L^2_x} = \|\nabla\tilde{\zeta}_b\|_{L^2_x} \lesssim \Gamma_b^{1-C\eta}$, which implies, by choosing $a \gg C\eta$, that

$$\|\nabla I_{N\lambda}\epsilon - \nabla I_{N\lambda}\tilde{\epsilon}\|_2 \lesssim \Gamma_b^{1+Ca}; \tag{6.25}$$

thus Lemma 6.7 implies in particular that

$$\partial_s \int \Psi_A|\epsilon|^2 \geq c_3 b \int_{A\leq|x|\leq 2A} |\epsilon|^2 - \Gamma_b^{1+Ca} - \Gamma_b^{a/2} \int |\nabla I_{N\lambda}\tilde{\epsilon}|^2 - \mathcal{F}(s) - \partial_s \mathcal{H}(s). \tag{6.26}$$

Proof of Lemma 6.7. Recall that since $\tilde{Q}_b$ is supported in $|x| \lesssim 1/b$, one has

$$\Psi_A|\tilde{Q}_b|^2 \equiv 0. \tag{6.27}$$

Thus

$$\Psi_A(|I_{N\lambda}\epsilon|^2 - |I_{N\lambda}\epsilon + \tilde{Q}_b|^2) = 0. \tag{6.28}$$

Now note

$$\begin{aligned} \Psi_A|\epsilon|^2 &= \Psi_A(|\epsilon|^2 - |I_{N\lambda}\epsilon|^2) \\ &\quad + \Psi_A(|I_{N\lambda}\epsilon|^2 - |I_{N\lambda}\epsilon + \tilde{Q}_b|^2) \\ &\quad + \Psi_A(-|I_{N\lambda}(\epsilon + \tilde{Q}_b)|^2 + |I_{N\lambda}\epsilon + \tilde{Q}_b|^2) \\ &\quad + \Psi_A(|I_{N\lambda}\epsilon + \tilde{Q}_b|^2). \end{aligned} \tag{6.29}$$

Observe that the second line of (6.29) is zero thanks to (6.27). Let

$$\mathcal{H} = \Psi_A(|\epsilon|^2 - |I_{N\lambda}\epsilon|^2) + \Psi_A(-|I_{N\lambda}(\epsilon + \tilde{Q}_b)|^2 + |I_{N\lambda}\epsilon + \tilde{Q}_b|^2). \tag{6.30}$$

Then $\mathcal{H}$ satisfies (6.24) since ϵ is bounded in H^s due to the bootstrap assumption (3.32), $\tilde{Q}_b$ is a nice function (uniformly in b), and $I_{N\lambda} - \mathrm{Id}$ removes all frequencies above $N(t)\lambda(t) \sim \lambda(t)^{-\delta}$ for some $\delta > 0$.

Moreover, we have

$$\frac{d}{ds}\int \Psi\left(\frac{x - x(t)}{A\lambda(t)}\right)|\epsilon|^2 = \frac{d}{ds}\frac{x - x(t)}{A(t)\lambda(t)}|I_{N\lambda}a|^2 - \partial_s\mathcal{H}, \tag{6.31}$$

where we recall the ansatz for a given by

$$a = \frac{1}{\lambda(t)}(\tilde{Q}_b + \epsilon)\left(\frac{x - x(t)}{\lambda(t)}\right)e^{i\gamma(t)}$$

and that the scaling in λ is L^2_x invariant.

As mentioned above, the role a plays for us is the same as the role played by u in the proof of [20, Lemma 4.5]. One may follow the computations leading to [20, (4.27)] and the formula above (4.27) in [20] to derive[24]

$$\begin{aligned}\frac{d}{ds}\int \Psi\Big(\frac{x-x(t)}{A\lambda(t)}\Big)|I_{N\lambda}a|^2 &\geq c_3 b\int_{A\leq|x|\leq 2A}|\epsilon^2| - \Gamma_b^{1+Ca} - \Gamma_b^{a/2}\int |I_{N\lambda}\nabla\epsilon|^2\\ &- \mathcal{F}_1(s) + 2\lambda^2\Im\int \Psi\Big(\frac{x-x(t)}{A\lambda(t)}\Big)\overline{I_N a}\,[I_N(a|a|^2) - I_N a|I_N a|^2]\\ &- \lambda^2\Big|\Psi\Big(\frac{x-x(t)}{A\lambda(t)}\Big)\tilde{I}_N a\,\overline{I_N a}\Big|\\ &+ 2\lambda^2\Im\int \Psi\Big(\frac{x-x(t)}{A\lambda(t)}\Big)\overline{I_N a}\,\big[I_N\big(|a+F|^2(a+F)\big) - I_N(|a|^2a)\big] \end{aligned} \tag{6.32}$$

where $\mathcal{F}_1$ satisfies (6.11) and consequently may be absorbed into $\mathcal{F}$. It has been explained in detail in [20] why the second line and third line of (6.32) can also be absorbed into $\mathcal{F}$. Note that in the final line of (6.32), we have used the fact that a satisfies a forced NLS (3.26).

We claim that for all $s_1 \in [s_0, s_+]$, one has

$$\int_{s_1}^{s_+}\Big|2\lambda^2\Im\int \Psi\Big(\frac{x-x(t)}{A\lambda(t)}\Big)\overline{I_N a}\,\big[I_N\big(|a+F|^2(a+F)\big) - I_N(|a|^2a)\big]\Big| \lesssim \lambda(s_1)^{\alpha_2} \tag{6.33}$$

for some $\alpha_2 > 0$. This is again similar to (6.13), since if one lets $t(s_1) \in [t_{k_1}, t_{k_1+1})$ and $T_+ = t(s_+)$, then using $\lambda^2 ds = dt$, one can bound the LHS via

$$\int_{t_{k_1}}^{T_+} (|a|_{L_x^4}^3 + |F|_{L_x^4}^3)\|F\|_{L_x^4}, \tag{6.34}$$

and proceed similarly to the proof of (6.13). We leave the details to the readers. ■

Step 5. In this step, we use the mass "conservation" law to combine Lemmas 6.5 and 6.7, and derive Lyapunov type control. For this part, we mostly directly refer to [20, 40]. It should be noted, however, that unlike the H^1 case from [40], or the $H^s, s > 0$, case from [20], here we will need to handle an almost conservation law rather than the exact conservation law. Indeed, if a solved the NLS, one would have

$$\frac{d}{ds}\|\tilde{Q}_b + \epsilon\|_2^2 = \frac{d}{ds}\|a\|_{L_x^2} \equiv 0.$$

In our case, a only solves a forced NLS (3.26) and thus does not enjoy precise mass conservation. We instead claim the following.

[24] In the original [20, (4.27)], there should be a $2\lambda^2$ before the $\Im(\cdots)$ term. Additionally, on the LHS of [20, (4.27)], Ψ_A should read $\Psi_{A\lambda}(x - x(t))$.

Lemma 6.9. *For all $s \in [s_0, s_+)$, one has*

$$\frac{d}{ds}\|\tilde{Q}_b + \epsilon\|_{L_x^2}^2 = \partial_s \mathcal{G}, \tag{6.35}$$

where $|\mathcal{G}(s)| \lesssim \lambda(s)^{\alpha_2}$ for some $\alpha_2 > 0$.

Proof. We compute

$$\frac{d}{ds}\|\tilde{Q}_b + \epsilon\|_{L_x^2}^2 = \frac{d}{ds}\|a\|_{L_x^2}^2 = \lambda^2 \frac{d}{dt}\|a\|_{L_x^2}^2, \tag{6.36}$$

which is bounded by

$$\lambda^2 \left| \int a\big(|a+F|^2(a+F) - |a|^2 a\big) \right|. \tag{6.37}$$

We need only verify that

$$\begin{aligned} &\int_s^{s_+} \lambda^2 \left| \int a\big(|a+F|^2(a+F) - |a|^2 a\big)\, ds \right| \\ &\qquad = \int_{t(s)}^{T_+} \left| \int a\big(|a+F|^2(a+F) - |a|^2 a\big)\, dt \right| \lesssim \lambda(s)^{\alpha_2} \end{aligned} \tag{6.38}$$

for some $\alpha_2 > 0$. This is again similar to (6.13) and we omit further details. ■

Thus, by expanding (6.35) and observing that Q does not depend on s, one has

$$\frac{d}{ds}\big(\|\tilde{Q}_b\|_{L_x^2}^2 - \|Q\|_{L_x^2}^2 + \|\epsilon\|_2^2 + 2(\epsilon_1, \Sigma_b) + 2(\epsilon_2, \Theta_2)\big) = \partial_s \mathcal{G}. \tag{6.39}$$

Now, we are ready to follow the computation in the proof of [40, Proposition 4]. We combine (6.23) and (6.21) with the help of (6.35), and, as in [40], we obtain

$$\begin{aligned} \frac{1}{100} c_3 \delta b f_1(s) + \frac{d}{ds}\int \Psi_A |\epsilon|^2 \geq \frac{1}{200} c_3 \delta \Xi(s) + \frac{1}{200} c_3 \delta b \left| \int I_{N\lambda} \nabla \epsilon \right|^2 \\ + \frac{1}{200} c_3 \delta b \int |\epsilon|^2 e^{-|y|} + \frac{1}{200} c_3 \delta b \Gamma_b - \mathcal{F} - \partial_s \mathcal{H}. \end{aligned}$$

We substitute

$$\frac{d}{ds}\int \Psi_A |\epsilon|^2 = -\big(\|\tilde{Q}_b\|_{L_x^2}^2 - \|Q\|_{L_x^2}^2 + \|\epsilon\|_2^2 + 2(\epsilon_1, \Sigma_b) + 2(\epsilon_2, \Theta_2)\big) - \partial_s \mathcal{G}, \tag{6.40}$$

and let $\mathcal{J}$ be defined as

$$\begin{aligned} \mathcal{J}(s) = \int |\tilde{Q}_b|^2 - \int |Q|^2 + 2(\epsilon_1, \Sigma) + 2(\epsilon_2, \Theta) + \int (1 - \Psi_A \epsilon^2) \\ - \frac{1}{100} c_3 \delta \big(\tilde{f}_1(b) - \int_0^b \tilde{f}_1(v)\, dv + b\{(\epsilon_2, \Lambda \Re \tilde{\zeta}_b) - (\epsilon_1, \Lambda \Im \tilde{\zeta}_b)\} - \mathcal{H} - \mathcal{G}, \end{aligned} \tag{6.41}$$

where

$$\tilde{f}_1(b) = \frac{b}{4}\|y\widetilde{Q}_b\|_2^2 + \frac{1}{2}\Im \int y\nabla\tilde{\zeta}_b\overline{\tilde{\zeta}_b}.$$

We then obtain

$$\partial_s \mathcal{J} \leq -Cb\left[\Gamma_b + \Xi + \int |\nabla I_{N\lambda}\tilde{\epsilon}|^2 + \int |\epsilon(s)|^2 e^{-|y|} + \int_{A\leq|x|\leq 2A} |\epsilon|^2\right] + \mathcal{F}, \qquad (6.42)$$

which corresponds to [20, (4.28)], though the definition of $\mathcal{J}$ now involves the corrections $\mathcal{H}$ and $\mathcal{G}$.

The main simple observation is that $\mathcal{J}$ is of size b^2, and the two extra correction terms are of size $\ll \Gamma_b^{100}$ and can be neglected. Hence, $\mathcal{J}$ with the extra corrections $\mathcal{H}$ and $\mathcal{G}$ can still serve as a Lyapunov function as in [20].

Estimates (6.21) and (6.42) ensure the dynamics remains in the log-log blowup regime, and are enough to close the bootstrap lemma 3.6. Indeed, the rest of the proof of the bootstrap lemma goes almost line for line as in [20], following the original scheme in [45]. We go over its proof quickly:

- One applies mass (almost) conservation law, Lemma 6.9, to upgrade (3.29) to (3.34). (In [20], one can just apply the exact mass conservation law.)
- One uses the monotonicity of $\mathcal{J}$ to upgrade (3.32) to (3.37).
- Estimate (6.9) implies in average sense $\lambda_s/\lambda \sim -b$, which is already enough to upgrade (3.30) to (3.30).
- Now the dynamics of λ is dictated by the dynamics of b, and b_s is governed by (1.14) and (6.42). This allows one to upgrade (3.31) to (3.36), and to upgrade (3.33) to (3.38).

This concludes the proof of the bootstrap lemma 3.6.

Finally, we may prove the main theorem:

Proof of Theorem 3.1. From the probabilistic local well-posedness of Lemma 4.2 and the energy estimates of Proposition 5.1, as we have detailed above, one obtains an exception set of small probability so that the bootstrap lemma 3.6 holds. As mentioned in the last two steps above, the dynamics of λ is dictated by that of b, whose dynamics is controlled by (6.42) and (1.14). This is sufficient to prove $\lambda(t)$ goes to zero at the desired rate; see [20, 40] for more details. This concludes the proof of Theorem 3.1. ∎

Acknowledgments. The authors thank Carlos Kenig and Gigliola Staffilani for interesting and helpful discussions. We also want to thank the anonymous referee for helpful comments and suggestions which have improved the manuscript.

Funding. C. Fan gratefully acknowledges support from a Simons Travel Grant and a start up funding from AMSS and NSFC grant No. 11688101. D. Mendelson gratefully acknowledges support from NSF grant DMS-1800697.

References

[1] Bényi, Á., Oh, T., Pocovnicu, O.: 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) MR 3350022

[2] Bényi, Á., Oh, T., Pocovnicu, O.: Wiener randomization on unbounded domains and an application to almost sure well-posedness of NLS. In: Excursions in Harmonic Analysis, Appl. Numer. Harmon. Anal. 4, Birkhäuser/Springer, 3–25 (2015) Zbl 1375.42040 MR 3411090

[3] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation. Geom. Funct. Anal. **3**, 209–262 (1993) Zbl 0787.35098 MR 1215780

[4] Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. Comm. Math. Phys. **166**, 1–26 (1994) Zbl 0822.35126 MR 1309539

[5] Bourgain, J.: Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. Comm. Math. Phys. **176**, 421–445 (1996) Zbl 0852.35131 MR 1374420

[6] Bourgain, J.: Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity. Int. Math. Res. Notices **1998**, 253–283 Zbl 0917.35126 MR 1616917

[7] Bourgain, J.: Global wellposedness of defocusing critical nonlinear Schrödinger equation in the radial case. J. Amer. Math. Soc. **12**, 145–171 (1999) Zbl 0958.35126 MR 1626257

[8] Bourgain, J., Bulut, A.: Invariant Gibbs measure evolution for the radial nonlinear wave equation on the 3d ball. J. Funct. Anal. **266**, 2319–2340 (2014) Zbl 1295.35323 MR 3150162

[9] Bringmann, B.: Almost-sure scattering for the radial energy-critical nonlinear wave equation in three dimensions. Anal. PDE **13**, 1011–1050 (2020) Zbl 1442.35259 MR 4109898

[10] Bringmann, B.: Stable blowup for the focusing energy critical nonlinear wave equation under random perturbations. Comm. Partial Differential Equations **45**, 1755–1777 (2020) Zbl 1459.35049 MR 4176915

[11] Bringmann, B.: Almost sure local well-posedness for a derivative nonlinear wave equation. Int. Math. Res. Notices **2021**, 8657–8697 Zbl 1473.35361 MR 4266148

[12] Brydges, D. C., Slade, G.: Statistical mechanics of the 2-dimensional focusing nonlinear Schrödinger equation. Comm. Math. Phys. **182**, 485–504 (1996) Zbl 0867.35090 MR 1447302

[13] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. I. Local theory. Invent. Math. **173**, 449–475 (2008) Zbl 1156.35062 MR 2425133

[14] Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. II. A global existence result. Invent. Math. **173**, 477–496 (2008) Zbl 1187.35233 MR 2425134

[15] Burq, N., Tzvetkov, N.: Probabilistic well-posedness for the cubic wave equation. J. Eur. Math. Soc. **16**, 1–30 (2014) Zbl 1295.35387 MR 3141727

[16] Cazenave, T.: Semilinear Schrödinger Equations. Courant Lecture Notes in Math. 10, New York Univ., Courant Institute of Mathematical Sciences, New York; Amer. Math. Soc., Providence, RI (2003) Zbl 1055.35003 MR 2002047

[17] Cazenave, T., Weissler, F. B.: Some remarks on the nonlinear Schrödinger equation in the critical case. In: Nonlinear Semigroups, Partial Differential Equations and Attractors (Washington, DC, 1987), Lecture Notes in Math. 1394, Springer, Berlin, 18–29 (1989) Zbl 0694.35170 MR 1021011

[18] Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T.: Almost conservation laws and global rough solutions to a nonlinear Schrödinger equation. Math. Res. Lett. **9**, 659–682 (2002) Zbl 1152.35491 MR 1906069

[19] Colliander, J., Oh, T.: Almost sure well-posedness of the cubic nonlinear Schrödinger equation below $L^2(\mathbb{T})$. Duke Math. J. **161**, 367–414 (2012) Zbl 1260.35199 MR 2881226

[20] Colliander, J., Raphaël, P.: Rough blowup solutions to the L^2 critical NLS. Math. Ann. **345**, 307–366 (2009) Zbl 1387.35547 MR 2529478

[21] Deng, Y., Nahmod, A. R., Yue, H.: Random tensors, propagation of randomness, and nonlinear dispersive equations. Invent. Math. **228**, 539–686 (2022) Zbl 07514024 MR 4411729

[22] Dodson, B.: Global well-posedness and scattering for the mass critical nonlinear Schrödinger equation with mass below the mass of the ground state. Adv. Math. **285**, 1589–1618 (2015) Zbl 1331.35316 MR 3406535

[23] Dodson, B., Lührmann, J., Mendelson, D.: Almost sure scattering for the 4D energy-critical defocusing nonlinear wave equation with radial data. Amer. J. Math. **142**, 475–504 (2020) MR 4084161

[24] Erdős, P.: Graph theory and probability. Canad. J. Math. **11**, 34–38 (1959) Zbl 0084.39602 MR 102081

[25] Glassey, R. T.: On the blowing up of solutions to the Cauchy problem for nonlinear Schrödinger equations. J. Math. Phys. **18**, 1794–1797 (1977) Zbl 0372.35009 MR 460850

[26] Glimm, J., Jaffe, A.: Quantum Physics. 2nd ed., Springer, New York (1987) Zbl 0461.46051 MR 887102

[27] Hadac, M., Herr, S., Koch, H.: Well-posedness and scattering for the KP-II equation in a critical space. Ann. Inst. H. Poincaré C Anal. Non Linéaire **26**, 917–941 (2009) Zbl 1169.35372 MR 2526409

[28] Keel, M., Tao, T.: Endpoint Strichartz estimates. Amer. J. Math. **120**, 955–980 (1998) Zbl 0922.35028 MR 1646048

[29] Kenig, C., Mendelson, D.: The focusing energy-critical nonlinear wave equation with random initial data. Int. Math. Res. Notices **2021**, 14508–14615 Zbl 1481.35282 MR 4324723

[30] Killip, R., Murphy, J., Visan, M.: Almost sure scattering for the energy-critical NLS with radial data below $H^1(\mathbb{R}^4)$. Comm. Partial Differential Equations **44**, 51–71 (2019) Zbl 1416.35241 MR 3933623

[31] Landman, M. J., Papanicolaou, G. C., Sulem, C., Sulem, P.-L.: Rate of blowup for solutions of the nonlinear Schrödinger equation at critical dimension. Phys. Rev. A (3) **38**, 3837–3843 (1988) MR 966356

[32] Lebowitz, J. L., Rose, H. A., Speer, E. R.: Statistical mechanics of the nonlinear Schrödinger equation. J. Statist. Phys. **50**, 657–687 (1988) Zbl 0925.35142 MR 939505

[33] Lührmann, J., Mendelson, D.: Random data Cauchy theory for nonlinear wave equations of power-type on $\mathbb{R}^3$. Comm. Partial Differential Equations **39**, 2262–2283 (2014) Zbl 1304.35788 MR 3259556

[34] Lührmann, J., Mendelson, D.: 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) Zbl 1408.35099 MR 3484682

[35] Merle, F.: Determination of blow-up solutions with minimal mass for nonlinear Schrödinger equations with critical power. Duke Math. J. **69**, 427–454 (1993) Zbl 0808.35141 MR 1203233

[36] Merle, F., Raphaël, P.: Sharp upper bound on the blow-up rate for the critical nonlinear Schrödinger equation. Geom. Funct. Anal. **13**, 591–642 (2003) Zbl 1061.35135 MR 1995801

[37] Merle, F., Raphaël, P.: On universality of blow-up profile for L^2 critical nonlinear Schrödinger equation. Invent. Math. **156**, 565–672 (2004) Zbl 1067.35110 MR 2061329

[38] Merle, F., Raphaël, P.: The blow-up dynamic and upper bound on the blow-up rate for critical nonlinear Schrödinger equation. Ann. of Math. (2) **161**, 157–222 (2005) Zbl 1185.35263 MR 2150386

[39] Merle, F., Raphaël, P.: Profiles and quantization of the blow up mass for critical nonlinear Schrödinger equation. Comm. Math. Phys. **253**, 675–704 (2005) Zbl 1062.35137 MR 2116733

[40] Merle, F., Raphaël, P.: On a sharp lower bound on the blow-up rate for the L^2 critical nonlinear Schrödinger equation. J. Amer. Math. Soc. **19**, 37–90 (2006) Zbl 1075.35077 MR 2169042

[41] Nahmod, A. R., Staffilani, G.: Almost sure well-posedness for the periodic 3D quintic nonlinear Schrödinger equation below the energy space. J. Eur. Math. Soc. **17**, 1687–1759 (2015) Zbl 1326.35353 MR 3361727

[42] Oh, T., Pocovnicu, O.: Probabilistic global well-posedness of the energy-critical defocusing quintic nonlinear wave equation on $\mathbb{R}^3$. J. Math. Pures Appl. (9) **105**, 342–366 (2016) Zbl 1343.35167 MR 3465807

[43] Oh, T., Sosoe, P., Tolomeo, L.: Optimal integrability threshold for Gibbs measures associated with focusing NLS on the torus. Invent. Math. **227**, 1323–1429 (2022) Zbl 07495371 MR 4384196

[44] Perelman, G.: On the blow up phenomenon for the critical nonlinear Schrödinger equation in 1D. In: Nonlinear Dynamics and Renormalization Group (Montreal, QC, 1999), CRM Proc. Lecture Notes 27, Amer. Math. Soc., Providence, RI, 147–164 (2001) Zbl 1055.35116 MR 1826598

[45] Planchon, F., Raphaël, P.: Existence and stability of the log-log blow-up dynamics for the L^2-critical nonlinear Schrödinger equation in a domain. Ann. Henri Poincaré **8**, 1177–1219 (2007) Zbl 1127.35062 MR 2355345

[46] Pocovnicu, O.: 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. **19**, 2521–2575 (2017) Zbl 1375.35278 MR 3668066

[47] Raphaël, P.: Stability of the log-log bound for blow up solutions to the critical non linear Schrödinger equation. Math. Ann. **331**, 577–609 (2005) Zbl 1082.35143 MR 2122541

[48] Tao, T.: Nonlinear Dispersive Equations. CBMS Reg. Conf. Ser. Math. 106, Amer. Math. Soc., Providence, RI (2006) Zbl 1106.35001 MR 2233925

[49] Thomann, L., Tzvetkov, N.: Gibbs measure for the periodic derivative nonlinear Schrödinger equation. Nonlinearity **23**, 2771–2791 (2010) Zbl 1204.35154 MR 2727169

[50] Tzvetkov, N.: Construction of a Gibbs measure associated to the periodic Benjamin–Ono equation. Probab. Theory Related Fields **146**, 481–514 (2010) Zbl 1188.35183 MR 2574736

[51] Weinstein, M. I.: Nonlinear Schrödinger equations and sharp interpolation estimates. Comm. Math. Phys. **87**, 567–576 (1982/83) Zbl 0527.35023 MR 691044

[52] Zhang, T., Fang, D.: Random data Cauchy theory for the generalized incompressible Navier–Stokes equations. J. Math. Fluid Mech. **14**, 311–324 (2012) Zbl 1294.35076 MR 2925111