# INVARIANT GIBBS MEASURES FOR THE 2-$d$ DEFOCUSING NONLINEAR WAVE EQUATIONS

TADAHIRO OH AND LAURENT THOMANN

Abstract. We consider the defocusing nonlinear wave equations (NLW) on the two-dimensional torus. In particular, we construct invariant Gibbs measures for the renormalized so-called Wick ordered NLW. We then prove weak universality of the Wick ordered NLW, showing that the Wick ordered NLW naturally appears as a suitable scaling limit of non-renormalized NLW with Gaussian random initial data.

Résumé. On considère les équations des ondes non-linéaires défocalisantes sur le tore de dimension deux. On construit des mesures de Gibbs invariantes pour les équations renormalisées au sens de Wick. On prouve ensuite une propriété d'universalité faible pour ces équations renormalisées, en montrant qu'elle apparaissent comme limites d'équations d'ondes non renormalisées avec conditions initiales aléatoires de loi gaussienne.

## 1. Introduction

1.1. **Nonlinear wave equations.** We consider the defocusing nonlinear wave equations (NLW) in two spatial dimensions:

$$\begin{cases} \partial_t^2 u - \Delta u + \rho u + u^{2m+1} = 0 \\ (u, \partial_t u)|_{t=0} = (\phi_0, \phi_1), \end{cases} \qquad (t,x) \in \mathbb{R} \times \mathcal{M}, \tag{1.1}$$

where $\rho \geq 0$ and $m \in \mathbb{N}$. When $\rho > 0$, (1.1) is also referred to as the nonlinear Klein-Gordon equation. We, however, simply refer to (1.1) as NLW and moreover restrict our attention to the real-valued setting. In the following, we mainly consider (1.1) on the two-dimensional torus $\mathcal{M} = \mathbb{T}^2 = (\mathbb{R}/\mathbb{Z})^2$ but we also provide a brief discussion when $\mathcal{M}$ is a two-dimensional compact Riemannian manifold without boundary or a bounded domain in $\mathbb{R}^2$ (with the Dirichlet or Neumann boundary condition). See Theorem 1.7 below.

Our main goal in this paper is to construct an invariant Gibbs measure for a renormalized version of (1.1) by studying dynamical properties of the renormalized equation.

1.2. **Gibbs measures and Wick renormalization.** With $v = \partial_t u$, we can write the equation (1.1) in the following Hamiltonian formulation:

$$\partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \frac{\partial H}{\partial(u,v)},$$

where $H = H(u,v)$ is the Hamiltonian given by

$$H(u,v) = \frac{1}{2}\int_{\mathbb{T}^2} \big(\rho u^2 + |\nabla u|^2\big)\, dx + \frac{1}{2}\int_{\mathbb{T}^2} v^2 dx + \frac{1}{2m+2}\int_{\mathbb{T}^2} u^{2m+2} dx. \tag{1.2}$$

2010 *Mathematics Subject Classification.* 35L71.

*Key words and phrases.* nonlinear wave equation; nonlinear Klein-Gordon equation; Gibbs measure; Wick ordering; Hermite polynomial; white noise functional; weak universality.

By drawing an analogy to the finite dimensional setting, the Hamiltonian structure of the equation and the conservation of the Hamiltonian suggest that the Gibbs measure $P_2^{(2m+2)}$ of the form: [1]

$$\text{“}dP_2^{(2m+2)} = Z^{-1}\exp(-\beta H(u,v))du\otimes dv\text{”} \tag{1.3}$$

is invariant under the dynamics of (1.1). [2] With (1.2), we can rewrite the formal expression (1.3) as

$$\begin{aligned} dP_2^{(2m+2)} &= Z^{-1}e^{-\frac{1}{2m+2}\int u^{2m+2}dx}e^{-\frac12\int(\rho u^2+|\nabla u|^2)dx}du\otimes e^{-\frac12\int v^2dx}dv\\ &\sim e^{-\frac{1}{2m+2}\int u^{2m+2}dx}d\mu, \end{aligned} \tag{1.4}$$

where $\mu$ is the Gaussian measure $\mu$ on $\mathcal{D}'(\mathbb{T}^2)\times\mathcal{D}'(\mathbb{T}^2)$ with the density [3]

$$d\mu = Z^{-1}e^{-\frac12\int(\rho u^2+|\nabla u|^2)dx}du\otimes e^{-\frac12\int v^2dx}dv. \tag{1.5}$$

Note that $\mu$ has a tensorial structure: $\mu=\mu_0\otimes\mu_1$, where the marginal measures $\mu_0$ and $\mu_1$ are given by

$$d\mu_0 = Z_0^{-1}e^{-\frac12\int(\rho u^2+|\nabla u|^2)dx}du \qquad\text{and}\qquad d\mu_1 = Z_1^{-1}e^{-\frac12\int v^2dx}dv. \tag{1.6}$$

Namely, $\mu_0$ is the Ornstein-Uhlenbeck measure and $\mu_1$ is the white noise measure on $\mathbb{T}^2$.

Recall that $\mu$ is the induced probability measure under the map: [4]

$$\omega\in\Omega\longmapsto(u,v)=\bigg(\sum_{n\in\mathbb{Z}^2}\frac{g_{0,n}(\omega)}{\langle n\rangle_\rho}e^{in\cdot x},\sum_{n\in\mathbb{Z}^2}g_{1,n}(\omega)e^{in\cdot x}\bigg), \tag{1.7}$$

where $\langle n\rangle_\rho = \sqrt{\rho+|n|^2}$ and $\{g_{0,n},g_{1,n}\}_{n\in\mathbb{Z}^2}$ is a sequence of independent standard complex-valued Gaussian random variables on a probability space $(\Omega,\mathcal{F},P)$ conditioned that $g_{j,-n}=\overline{g_{j,n}}$, $n\in\mathbb{Z}^2$, $j=0,1$. In view of (1.7), it is easy to see that $\mu$ is supported on

$$\mathcal{H}^s(\mathbb{T}^2) := H^s(\mathbb{T}^2)\times H^{s-1}(\mathbb{T}^2), \quad s<0.$$

Moreover, we have $\mu(\mathcal{H}^0(\mathbb{T}^2))=0$. This implies that $\int u^{2m+2}dx=\infty$ almost surely with respect to $\mu$. In particular, the right-hand side of (1.4) would not be a probability measure, thus requiring a *renormalization* of the potential part of the Hamiltonian. In the two-dimensional case, it is known that a Wick ordering suffices for this purpose. See Simon [31] and Glimm-Jaffe [15]. Also, see Da Prato-Tubaro [12] for a concise discussion on $\mathbb{T}^2$, where the Gibbs measures naturally appear in the context of the stochastic quantization equation.

In the following, we give a brief review of the Wick renormalization on $\mathbb{T}^2$. See [12] for more details. Let $u$ denote a typical element under $\mu_0$ defined in (1.6). Since $u\notin L^2(\mathbb{T}^2)$ almost surely, we have

$$\int_{\mathbb{T}^2}u^2dx = \lim_{N\to\infty}\int_{\mathbb{T}^2}(\mathbf{P}_Nu)^2dx=\infty$$

almost surely, where $\mathbf{P}_N$ is the Dirichlet projection onto the frequencies $\{|n|\le N\}$.

1. Henceforth, we use $Z$, $Z_N$, etc. to denote various normalizing constants so that the corresponding measures are probability measures when appropriate.

2. We simply set $\beta=1$ in the following. While our analysis holds for any $\beta>0$, the resulting (renormalized) Gibbs measures are mutually singular for different values of $\beta>0$. See [28].

3. On $\mathbb{T}^2$, we need to assume $\rho>0$ in order to avoid a problem at the zeroth frequency. See (1.7) below. In the case of a bounded domain in $\mathbb{R}^2$ with the Dirichlet boundary condition, we can take $\rho=0$.

4. We drop the harmless factor $2\pi$ in the following.

For each $x \in \mathbb{T}^2$, $\mathbf{P}_N u(x)$ is a mean-zero real-valued Gaussian random variable with variance [5]

$$\sigma_N \stackrel{\text{def}}{=} \mathbb{E}[(\mathbf{P}_N u)^2(x)] = \sum_{|n| \le N} \frac{1}{\rho + |n|^2} \sim \log N. \tag{1.8}$$

This motivates us to define the Wick ordered monomial $:(\mathbf{P}_N u)^k:$ by

$$:(\mathbf{P}_N u)^k(x): = H_k(\mathbf{P}_N u(x); \sigma_N) \tag{1.9}$$

in a pointwise manner. Here, $H_k(x;\sigma)$ is the Hermite polynomial of degree $k$ defined in (2.1). Then, with (1.7) and (1.8), it is easy to see that the random variables $X_N(u)$ defined by

$$X_N(u) = \int_{\mathbb{T}^2} :(\mathbf{P}_N u)^2(x): dx$$

have uniformly bounded second moments and converge to some random variable in $L^2(d\mu_0)$ which we denote by

$$X_\infty(u) = \int_{\mathbb{T}^2} :u^2: dx \in L^2(d\mu_0).$$

In view of the Wiener chaos estimate (Lemma 2.2), we see that $X_N(u)$ also converges to $X_\infty(u)$ in $L^p(d\mu_0)$, $p < \infty$.

In general, given any $m \in \mathbb{N}$, one can show that the limit

$$\int_{\mathbb{T}^2} :u^{2m+2}: dx = \lim_{N\to\infty} \int_{\mathbb{T}^2} :(\mathbf{P}_N u)^{2m+2}: dx \tag{1.10}$$

exists in $L^p(\mu)$ for any finite $p \ge 1$. Moreover, we have the following proposition.

**Proposition 1.1.** *Let $m \in \mathbb{N}$. Then, $R_N(u) \stackrel{\text{def}}{=} e^{-\frac{1}{2m+2}\int_{\mathbb{T}^2} :(\mathbf{P}_N u)^{2m+2}: dx} \in L^p(\mu)$ for any finite $p \ge 1$ with a uniform bound in $N$, depending on $p \ge 1$. Moreover, for any finite $p \ge 1$, $R_N(u)$ converges to some $R(u)$ in $L^p(\mu)$ as $N \to \infty$.*

This proposition follows from the hypercontractivity of the Ornstein-Uhlenbeck semigroup and Nelson's estimate [26]. See also [12, 30]. Denoting the limit $R(u) \in L^p(\mu)$ by

$$R(u) = e^{-\frac{1}{2m+2}\int_{\mathbb{T}^2} :u^{2m+2}: dx},$$

Proposition 1.1 allows us to define the Gibbs measure $P_2^{(2m+2)}$ associated with the Wick ordered Hamiltonian:

$$H_{\text{Wick}}(u,v) = \frac{1}{2}\int_{\mathbb{T}^2} \left(\rho u^2 + |\nabla u|^2\right) dx + \frac{1}{2}\int_{\mathbb{T}^2} v^2 dx + \frac{1}{2m+2}\int_{\mathbb{T}^2} :u^{2m+2}: dx$$

by

$$\begin{aligned} dP_2^{(2m+2)} &= Z^{-1} e^{-H_{\text{Wick}}(u,v)} du \otimes dv = Z^{-1} e^{-\frac{1}{2m+2}\int_{\mathbb{T}^2} :u^{2m+2}: dx} d\mu \\ &= Z^{-1} R(u) d\mu. \end{aligned}$$

5. Note that $\sigma_N$ defined in (1.8) is independent of $x \in \mathbb{T}^2$. When $\mathcal{M}$ is a two-dimensional compact Riemannian manifold without boundary or a bounded domain in $\mathbb{R}^2$, the variance $\sigma_N(x) = \mathbb{E}[(\mathbf{P}_N u)^2(x)]$ depends on $x \in \mathcal{M}$ but satisfies the logarithmic bound in $N$. See (1.12) below.

It follows from Proposition 1.1 that $P_2^{(2m+2)} \ll \mu$ and, in particular, $P_2^{(2m+2)}$ is a probability measure on $\mathcal{H}^s(\mathbb{T}^2) \setminus \mathcal{H}^0(\mathbb{T}^2)$, $s < 0$. Moreover, defining $P_{2,N}^{(2m+2)}$ by

$$dP_{2,N}^{(2m+2)} = Z_N^{-1} R_N(u) d\mu,$$

we see that $P_{2,N}^{(2m+2)}$ converges "uniformly" to $P_2^{(2m+2)}$ in the sense that given any $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that

$$\big|P_{2,N}^{(2m+2)}(A) - P_2^{(2m+2)}(A)\big| < \varepsilon$$

for any $N \geq N_0$ and any measurable set $A \subset \mathcal{H}^s(\mathbb{T})$, $s < 0$.

Lastly, let us briefly discuss the construction of the Gibbs measure $P_2^{(2m+2)}$ when $\mathcal{M}$ is a two-dimensional compact Riemannian manifold without boundary or a bounded domain in $\mathbb{R}^2$ (with the Dirichlet or Neumann boundary condition). In this case, the Gaussian measure $\mu$ in (1.5) represents the induced probability measure under the map:

$$\omega \in \Omega \longmapsto (u, v) = \bigg( \sum_{n \in \mathbb{N}} \frac{g_{0,n}(\omega)}{(\rho + \lambda_n^2)^{\frac{1}{2}}} \varphi_n(x), \sum_{n \in \mathbb{N}} g_{1,n}(\omega) \varphi_n(x) \bigg), \tag{1.11}$$

where $\{\varphi_n\}_{n \in \mathbb{N}}$ is an orthonormal basis of $L^2(\mathcal{M})$ consisting of eigenfunctions of the Laplace-Beltrami operator $-\Delta$ with the corresponding eigenvalues $\{\lambda_n^2\}_{n \in \mathbb{N}}$, which we assume to be arranged in the increasing order. It is easy to see from (1.11) that $\mu$ is supported on $\mathcal{H}^s(\mathcal{M}) \setminus \mathcal{H}^0(\mathcal{M})$, $s < 0$.

Given $N \in \mathbb{N}$, we define $\sigma_N$ by

$$\sigma_N(x) = \mathbb{E}[(\mathbf{P}_N u_N)^2(x)] = \sum_{\lambda_n \leq N} \frac{\varphi_n(x)^2}{\rho + \lambda_n^2} \lesssim \log N, \tag{1.12}$$

where $\mathbf{P}_N$ denotes the spectral projector defined by

$$\mathbf{P}_N u = \sum_{\lambda_n \leq N} \widehat{u}(n) \varphi_n.$$

Note that unlike the situation on $\mathbb{T}^2$, $\sigma_N(x)$ now depends on $x \in \mathcal{M}$. The last inequality in (1.12), however, holds independently of $x \in \mathcal{M}$ thanks to Weyl's law $\lambda_n \approx n^{\frac{1}{2}}$ (see [33, Chapter 14]) and [8, Proposition 8.1]. With this definition of $\sigma_N(x)$, we can define the Wick ordered monomials $:(\mathbf{P}_N u)^k:$ as in (1.9) and $:u^k:$ by the limiting procedure. Then, the discussion above for $\mathbb{T}^2$, in particular Proposition 1.1, also holds on $\mathcal{M}$. See Section 4 of [30]. While the presentation in [30] is given in the complex-valued setting, a straightforward modification yields the corresponding result for the real-valued setting.

In the next subsection, we discuss the dynamical problem. Our main goal in this paper is to construct dynamics for the renormalized equation associated with the Wick ordered Hamiltonian $H_{\text{Wick}}$ with initial data distributed according to the Gibbs measure $P_2^{(2m+2)}$.

1.3. **Dynamical problem: Wick ordered NLW.** We now consider the following dynamical problem on $\mathbb{T}^2$ associated with the Wick ordered Hamiltonian:

$$\begin{cases} \partial_t \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \dfrac{\partial H_{\text{Wick}}}{\partial(u, v)} \\ (u, v)|_{t=0} = (\phi_0^\omega, \phi_1^\omega), \end{cases} \tag{1.13}$$

where the initial data $(\phi_0^\omega, \phi_1^\omega)$ is distributed according to the Gibbs measure $P_2^{(2m+2)}$. In view of the absolute continuity of $P_2^{(2m+2)}$ with respect to the Gaussian measure $\mu$ (Proposition 1.1), we consider the random initial data $(\phi_0^\omega, \phi_1^\omega)$ distributed according to $\mu$ in the following discussion. Namely, we assume that

$$\big(\phi_0^\omega, \phi_1^\omega\big) = \bigg( \sum_{n\in\mathbb{Z}^2} \frac{g_{0,n}(\omega)}{\langle n\rangle_\rho} e^{in\cdot x}, \sum_{n\in\mathbb{Z}^2} g_{1,n}(\omega) e^{in\cdot x}\bigg), \tag{1.14}$$

where $\{g_{0,n}, g_{1,n}\}_{n\in\mathbb{Z}^2}$ is as in (1.7). Note that, at this point, the potential part $\frac{1}{2m+2}\int_{\mathbb{T}^2} :u^{2m+2}: dx$ of the Wick ordered Hamiltonian is defined only for $u$ distributed according to the Gaussian measure $\mu$ via (1.10). In the following, we extend this definition to a wider class of functions in order to treat the Cauchy problem (1.13).

Given $N\in\mathbb{N}$, define the truncated Wick ordered Hamiltonian $H^N_{\text{Wick}}$ by

$$H^N_{\text{Wick}}(u,v) = \frac12 \int_{\mathbb{T}^2} \big(\rho u^2 + |\nabla u|^2\big)\, dx + \frac12\int_{\mathbb{T}^2} v^2 dx + \frac{1}{2m+2}\int_{\mathbb{T}^2} :(\mathbf{P}_N u)^{2m+2}: dx \tag{1.15}$$

and consider the associated Hamiltonian dynamics:

$$\begin{cases} \partial_t \begin{pmatrix} u_N \\ v_N \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \dfrac{\partial H^N_{\text{Wick}}}{\partial(u_N, v_N)} \\ (u_N, v_N)|_{t=0} = (\phi_0^\omega, \phi_1^\omega). \end{cases}$$

Thanks to (1.9) and $\partial_x H_k(x;\sigma) = kH_{k-1}(x;\sigma)$, we can rewrite the system (1.15) as the following truncated Wick ordered NLW:

$$\begin{cases} \partial_t^2 u_N - \Delta u_N + \rho u_N + \mathbf{P}_N\big[\, :(\mathbf{P}_N u_N)^{2m+1}: \,\big] = 0 \\ (u_N, \partial_t u_N)|_{t=0} = (\phi_0^\omega, \phi_1^\omega), \end{cases} \tag{1.16}$$

where the truncated Wick ordered nonlinearity is interpreted as

$$\mathbf{P}_N\big[\, :(\mathbf{P}_N u_N)^{2m+1}: \,\big] = \mathbf{P}_N\big[H_{2m+1}(\mathbf{P}_N u_N; \sigma_N)\big].$$

Let $z = z^\omega$ denote the random linear solution:

$$z(t) = S(t)(\phi_0^\omega, \phi_1^\omega) = \cos(t\langle\nabla\rangle_\rho)\phi_0^\omega + \frac{\sin(t\langle\nabla\rangle_\rho)}{\langle\nabla\rangle_\rho}\phi_1^\omega, \tag{1.17}$$

where $\langle\nabla\rangle_\rho = \sqrt{\rho-\Delta}$. In view of the Duhamel formula, it is natural to decompose the solution $u_N$ to (1.16) as

$$u_N = z + w_N.$$

Note that we have $\mathbf{P}_N w_N = w_N$. By recalling the following identities for the Hermite polynomials:

$$H_k(x+y) = \sum_{\ell=0}^k \binom{k}{\ell} H_\ell(y)\cdot x^{k-\ell} \qquad \text{and} \qquad H_k(x;\sigma) = \sigma^{\frac{k}{2}} H_k(\sigma^{-\frac12}x), \tag{1.18}$$

we have

$$\begin{aligned} :(\mathbf{P}_N u_N)^{2m+1}: &= H_{2m+1}(z_N + w_N; \sigma_N) \\ &= \sum_{\ell=0}^{2m+1} \binom{2m+1}{\ell} H_\ell(z_N;\sigma_N)\cdot w_N^{2m+1-\ell}, \end{aligned} \tag{1.19}$$

where $z_N = \mathbf{P}_N z$. This shows that applying the Wick ordering to the monomial

$$(\mathbf{P}_N u_N)^{2m+1} = (z_N + w_N)^{2m+1} = \sum_{\ell=0}^{2m+1} \binom{2m+1}{\ell} z_N^\ell \cdot w_N^{2m+1-\ell} \tag{1.20}$$

is equivalent to Wick ordering all the monomials $z_N^\ell$. Namely, replacing each $z^\ell$ in (1.20) by

$$:\! z_N^\ell \!: = H_\ell(z_N; \sigma_N)$$

yields the Wick ordered monomial $:\!(\mathbf{P}_N u_N)^{2m+1}\!:$ via (1.19). In Proposition 2.3 below, we prove that

$$:\! z_N^\ell \!: \in L^p(\Omega; L^q([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2)))$$

for any $p, q, r < \infty$, $T > 0$, and $\varepsilon > 0$ with a bound uniform in $N$. Moreover, the sequence $\big\{ :\! z_N^\ell \!: \big\}_{N\in\mathbb{N}}$ is a Cauchy sequence in the same space, thus allowing us to define

$$:\! z^\ell \!: \, = \, :\! z_\infty^\ell \!: \stackrel{\text{def}}{=} \lim_{N\to\infty} :\! z_N^\ell \!: \tag{1.21}$$

in $L^p(\Omega; L^q([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2)))$ for any $p, q, r < \infty$, $T > 0$, and $\varepsilon > 0$ (and for any $\ell \in \mathbb{N}$). Now, consider a function $u$ of the form

$$u = z + w \tag{1.22}$$

for some "nice" $w$. Then, we can use (1.18) and (1.21) to define the Wick ordered monomial $:\! u^{2m+1} \!:$ for functions $u$ of the form (1.22) by

$$:\! u^{2m+1} \!: \, = \, :\!(z+w)^{2m+1}\!: = \sum_{\ell=0}^{2m+1} \binom{2m+1}{\ell} :\! z^\ell \!: \cdot\, w^{2m+1-\ell}. \tag{1.23}$$

Hence, we finally arrive at the *defocusing Wick ordered NLW*:

$$\begin{cases} \partial_t^2 u - \Delta u + \rho u + :\! u^{2m+1} \!: = 0 \\ (u, \partial_t u)|_{t=0} = (\phi_0^\omega, \phi_1^\omega), \end{cases} \tag{1.24}$$

where $(\phi_0^\omega, \phi_1^\omega)$ is as in (1.14).

Before we state our main result, we first recall two critical regularities associated with (1.1) on $\mathbb{R}^2$ with $\rho = 0$. On the one hand, the scaling symmetry for (1.1) induces the so-called scaling critical Sobolev index: $s_1 = 1 - \frac{1}{m}$. On the other hand, the Lorentzian invariance (conformal symmetry) induces another critical regularity: $s_2 = \frac{3}{4} - \frac{1}{2m}$ (at least in the focusing case). Hence, we set $s_{\text{crit}}$ by

$$s_{\text{crit}} = \max\bigg(1 - \frac{1}{m}, \frac{3}{4} - \frac{1}{2m}\bigg) = \begin{cases} \frac{1}{4} & \text{if} \quad m = 1, \\ 1 - \frac{1}{m} & \text{if} \quad m \geq 2. \end{cases}$$

We now state our main result.

**Theorem 1.2.** *Let $\mathcal{M} = \mathbb{T}^2$, $m \in \mathbb{N}$, and $\rho > 0$. Then, the Wick ordered NLW* (1.24) *is almost surely locally well-posed with respect to the Gaussian measure $\mu$ defined in* (1.5)*. More precisely, letting $(\phi_0^\omega, \phi_1^\omega)$ be as in* (1.14)*, there exist $C, c > 0$ such that for each $T \ll 1$, there exists a set $\Omega_T \subset \Omega$ with the following properties:*

(i) $P(\Omega_T^c) \leq C \exp\big(-\frac{1}{T^c}\big)$,

(ii) *For each $\omega \in \Omega_T$, there exists a (unique) solution $u$ to (1.24) with $(u, \partial_t u)|_{t=0} = (\phi_0^\omega, \phi_1^\omega)$ in the class*

$$S(t)(\phi_0^\omega, \phi_1^\omega) + C([-T,T]; H^s(\mathbb{T}^2)) \cap X_T^{s,\frac{1}{2}+} \subset C([-T,T]; H^{-\varepsilon}(\mathbb{T}^2))$$

*for any $s \in (s_{\text{crit}}, 1)$ and $\varepsilon > 0$. Here, $X_T^{s,\frac{1}{2}+}$ denotes the local-in-time version of the hyperbolic Sobolev space. See Section 3.*

We emphasize that the Wick ordered NLW (1.24) is defined *only* for functions $u$ of the form (1.22). Then, the residual term $w = u - z$ satisfies the following perturbed Wick ordered NLW:

$$\begin{cases} \partial_t^2 w - \Delta w + \rho w + :(w+z)^{2m+1}: = 0 \\ (w, \partial_t w)|_{t=0} = (0,0). \end{cases} \tag{1.25}$$

By writing (1.25) in the Duhamel formulation, we obtain

$$\begin{aligned} w(t) &= -\int_0^t \frac{\sin((t-t')\langle\nabla\rangle_\rho)}{\langle\nabla\rangle_\rho} :(w+z)^{2m+1}(t'): dt' \\ &= -\sum_{\ell=0}^{2m+1} \int_0^t \frac{\sin((t-t')\langle\nabla\rangle_\rho)}{\langle\nabla\rangle_\rho} \binom{2m+1}{\ell} :z^\ell(t'): \cdot w^{2m+1-\ell}(t') dt'. \end{aligned} \tag{1.26}$$

We prove Theorem 1.2 by solving the fixed point problem (1.26) for $w$ in $C([-T,T]; H^s(\mathbb{T}^2)) \cap X_T^{s,\frac{1}{2}+}$, $s > s_{\text{crit}}$. In Section 2, we study the regularity of the random linear solution $z$ and the associated Wick ordered monomials $:z^\ell:$. In particular, while they are rough, $:z^\ell:$ enjoys enhanced integrability both in space and time. See Proposition 2.3. In Section 3, we then use the standard Fourier restriction norm method to solve the fixed point problem (1.26). The original idea of this argument with the decomposition (1.22) appears in McKean [25] and Bourgain [4] in the context of the nonlinear Schrödinger equations on $\mathbb{T}^d$, $d = 1, 2$. See also Burq-Tzvetkov [9]. In the field of the stochastic PDEs, this method is known as Da Prato-Debussche trick [11].

**Remark 1.3.** As in the study of singular stochastic PDEs, our proof consists of factorizing the ill-defined solution map: $(\phi_0^\omega, \phi_1^\omega) \mapsto u$ into a canonical lift followed by a (continuous) solutions map $\Psi$:

$$\begin{aligned} (\phi_0^\omega, \phi_1^\omega) \xmapsto{\text{lift}} (z_1^\omega, z_3^\omega, \dots, z_{2m+1}^\omega) &\xmapsto{\Psi} w \in C([-T,T]; H^s(\mathbb{T}^2)) \\ &\longmapsto u = z + w \in C([-T,T]; H^{-\varepsilon}(\mathbb{T}^2)), \end{aligned}$$

for $s \in (s_{\text{crit}}, 1)$ and $\varepsilon > 0$, where $z_k \stackrel{\text{def}}{=} :z^k:$. On the one hand, we use probability theory to construct the data set $\{z_{2j+1}\}_{j=0}^m$ in the first step. On the other hand, the second step is entirely deterministic. Moreover, the solution map $\Psi$ in the second step is continuous from $\prod_{j=0}^m S_T^j$ to $X_T^{s,\frac{1}{2}+}$, where $S_T^j$ denotes some appropriate Strichartz space for $z_{2j+1}$. See Section 3.

**Remark 1.4.** The same almost sure local well-posedness holds for the truncated Wick ordered NLW (1.16). More precisely, we can choose $\Omega_T$, independent of $N \in \mathbb{N}$, such that the statement in Theorem 1.2 holds for (1.24) and (1.16). Moreover, by possibly shrinking the time, one can also prove that the solution $u_N = u_N^\omega$ to (1.16) converges to the solution $u = u^\omega$ to (1.24) as $N \to \infty$.

Once we have almost sure local well-posedness of (1.24), the invariant measure argument by Bourgain [3, 4] yields the following almost sure global well-posedness of (1.24) and invariance of the Gibbs measure $P_2^{(2m+2)}$.

**Theorem 1.5.** *Let $\mathcal{M} = \mathbb{T}^2$, $m \in \mathbb{N}$, and $\rho > 0$. Then, the defocusing Wick ordered NLW* (1.24) *is almost surely globally well-posed with respect to the Gibbs measure $P_2^{(2m+2)}$. Moreover, $P_2^{(2m+2)}$ is invariant under the dynamics of* (1.24).

The proof of Theorem 1.5 exploits the invariance of the truncated Gibbs measure $P_{2,N}^{(2m+2)}$ for the truncated Wick ordered NLW (1.16) and combines it with an approximation argument. See Remark 1.4. As this argument is standard by now, we omit the proof. See Bourgain [4] and Burq-Tzvetkov [10] for details.

**Remark 1.6.** We point that the convergence result in Remark 1.4 and invariance of the Gibbs measure in Theorem 1.5 already appear (without a proof) in the lecture note by Bourgain [5]. See [5, Theorem 111 on p. 63] and a comment that follows (118) on p. 64 in [5]. To the best of our knowledge, however, there seems to be no proof available in a published paper. In fact, one of the main purposes of this paper is to present the details of the proof of Bourgain's claim in [5].

Next, we briefly discuss the situation when the spatial domain $\mathcal{M}$ is a two-dimensional compact Riemannian manifold without boundary or a bounded domain in $\mathbb{R}^2$ (with the Dirichlet or Neumann boundary condition). In this case, one can exploit the invariance of the truncated Gibbs measures $P_{2,N}^{(2m+2)}$ for (1.16) to construct global-in-time *weak* solutions (without uniqueness) to the Wick ordered NLW (1.24). Moreover, it also allows us to establish invariance of the Gibbs measure $P_2^{(2m+2)}$ in some mild sense.

**Theorem 1.7.** *Let $m \in \mathbb{N}$ and $\rho > 0$. Let $\mathcal{M}$ be a two-dimensional compact Riemannian manifold without boundary or a bounded domain in $\mathbb{R}^2$ (with the Dirichlet or Neumann boundary condition). In the latter case with the Dirichlet boundary condition, we can also take $\rho = 0$. Then, there exists a set $\Sigma$ of full measure with respect to $P_2^{(2m+2)}$ such that for every $\phi \in \Sigma$, the defocusing Wick ordered NLW* (1.24) *with initial data distributed according to $P_2^{(2m+2)}$ has a global-in-time solution $u \in C(\mathbb{R}; H^s(\mathcal{M}))$ for any $s < 0$. Moreover, for all $t \in \mathbb{R}$, the law of the random function $(u, \partial_t u)(t)$ is given by $P_2^{(2m+2)}$.*

In [30], we proved an analogous result for the defocusing Wick ordered nonlinear Schrödinger equations on $\mathcal{M}$. Theorem 1.7 follows from repeating the argument presented in [30] with systematic modifications and thus we omit details. See also [1, 11, 8, 29]. The main ingredient for Theorem 1.7 is to establish tightness (= compactness) of measures $\nu_N$ on space-time functions, emanating from the truncated Gibbs measure $P_{2,N}^{(2m+2)}$ and then upgrading the weak convergence of $\nu_N$ (up to a subsequence) to an almost sure convergence of the corresponding random variables via Skorokhod's theorem. Due to the compactness argument, Theorem 1.7 claims only the existence of a global-in-time solution $u$. Lastly, note that Theorem 1.7 only claims that the law of the $\mathcal{H}^s$-valued random variable $(u, \partial_t u)(t)$ is given by the Gibbs measure $P_2^{(2m+2)}$ for any $t \in \mathbb{R}$. In particular, this mild invariance for a general geometric setting is weaker than the invariance stated in Theorem 1.5 for the Wick ordered NLW (1.24) on $\mathbb{T}^2$.

**Remark 1.8.** On the one hand, the defocusing/focusing nature of the equation does not play any role in the almost sure local well-posedness result (Theorem 1.2) and thus Theorem 1.2 also holds in the focusing case. It can also be extended to Wick ordered even power monomials in the equation. On the other hand, the defocusing nature of the equation plays a crucial role in the proof of Proposition 1.1 and hence in Theorems 1.5 and 1.7. In the focusing case (i.e. with $-u^{2m+1}$, $m \in \mathbb{N}$, in (1.1)), it is known that the Gibbs measure can not be normalized in the two dimensional case. See Brydges-Slade [6]. Lastly, we point out that in the case of the quadratic nonlinearity (which is neither defocusing nor focusing), one can introduce the following modified Gibbs measure:

$$dP_2^{(3)} = Z^{-1} e^{-\frac{1}{3}\int :u^3: - A(\int :u^2:)^2} d\mu$$

for sufficiently large $A \gg 1$ and study the associated dynamical problem. See [5] for the construction of this modified Gibbs measure $P_2^{(3)}$.

1.4. **Wick ordered NLW as a scaling limit.** As an application of the local well-posedness argument, we show how the Wick ordered NLW (1.24) appears as a scaling limit of non-renormalized NLW equations on dilated tori. This part of the discussion is strongly motivated by the weak universality result for the Wick ordered stochastic NLW on $\mathbb{T}^2$ studied by the first author with Gubinelli and Koch in [16].

Fix $\rho > 0$. Given small $\varepsilon > 0$, we consider the following non-renormalized NLW equation on a dilated torus $\mathbb{T}_\varepsilon^2 \stackrel{\text{def}}{=} (\varepsilon^{-1}\mathbb{T})^2$:

$$\begin{cases} \partial_t^2 v_\varepsilon - \Delta v_\varepsilon + \rho_\varepsilon v_\varepsilon = f(v_\varepsilon) \\ (v_\varepsilon, \partial_t v_\varepsilon)|_{t=0} = (\psi_{\varepsilon,0}^\omega, \psi_{\varepsilon,1}^\omega), \end{cases} \qquad (t,x) \in \mathbb{R} \times \mathbb{T}_\varepsilon^2 \tag{1.27}$$

with Gaussian random initial data $(\psi_{\varepsilon,0}^\omega, \psi_{\varepsilon,1}^\omega)$, where $f : \mathbb{R} \to \mathbb{R}$ is a smooth odd [6] function with the following bound:

$$|f^{(4)}(x)| \lesssim 1 + |x|^M \tag{1.28}$$

for some $M \geq 0$ and $\rho_\varepsilon$ is a parameter to be chosen later. In the following, we choose $\psi_{\varepsilon,0}^\omega$ and $\psi_{\varepsilon,1}^\omega$ to be a smoothed Ornstein-Uhlenbeck process and a smoothed white noise on $\mathbb{T}_\varepsilon^2$, respectively. For the sake of concreteness, we set [7]

$$(\psi_{\varepsilon,0}^\omega, \psi_{\varepsilon,1}^\omega) = \left( \sum_{\substack{n \in (\varepsilon\mathbb{Z})^2 \\ |n| \leq 1}} \frac{g_{0,\varepsilon^{-1}n}}{\varepsilon^{-1}\sqrt{\varepsilon^2\rho + |n|^2}} \frac{e^{in\cdot x}}{\varepsilon^{-1}}, \sum_{\substack{n \in (\varepsilon\mathbb{Z})^2 \\ |n| \leq 1}} g_{1,\varepsilon^{-1}n} \frac{e^{in\cdot x}}{\varepsilon^{-1}} \right),$$

6. It follows from the proof of Theorem 1.9 that it suffices to assume that $f(0) = f''(0) = 0$ for the cubic case considered in Theorem 1.9.

7. Note that $\{\varepsilon e^{in\cdot x}\}_{n \in (\varepsilon\mathbb{Z})^2}$ forms an orthonormal basis of $L^2(\mathbb{T}_\varepsilon^2)$. Moreover, recall that the Fourier-Wiener series

$$\sum_{n \in (\varepsilon\mathbb{Z})^2} \frac{g_{0,n}}{\varepsilon^{-1}|n|} \frac{e^{in\cdot x}}{\varepsilon^{-1}}$$

represents the periodic Wiener process on $\mathbb{T}_\varepsilon^2$.

where $\{g_{0,n}, g_{1,n}\}_{n\in\mathbb{Z}^2}$ is as in (1.7). Our main goal is to study the behavior of the solution to (1.27) as $\varepsilon \to 0$ by applying a suitable scaling.

Let $u_\varepsilon(t,x) \overset{\text{def}}{=} \varepsilon^{-1} v_\varepsilon(\varepsilon^{-1}x, \varepsilon^{-1}t)$. Then, $u_\varepsilon$ satisfies

$$\begin{cases} \partial_t^2 u_\varepsilon - \Delta u_\varepsilon + \rho u_\varepsilon = \varepsilon^{-3}\big\{ f(\varepsilon u_\varepsilon) + \varepsilon(\varepsilon^2\rho - \rho_\varepsilon)u_\varepsilon \big\} \\ (u_\varepsilon, \partial_t u_\varepsilon)|_{t=0} = (\phi^\omega_{\varepsilon,0}, \phi^\omega_{\varepsilon,1}), \end{cases} \qquad (t,x)\in\mathbb{R}\times\mathbb{T}^2, \tag{1.29}$$

where $(\phi^\omega_{\varepsilon,0}, \phi^\omega_{\varepsilon,1})$ on $\mathbb{T}^2$ is given by

$$(\phi^\omega_{\varepsilon,0}, \phi^\omega_{\varepsilon,1}) = \Bigg( \sum_{\substack{n\in\mathbb{Z}^2 \\ |n|\le\varepsilon^{-1}}} \frac{g_{0,n}}{\sqrt{\rho+|n|^2}} e^{in\cdot x}, \sum_{\substack{n\in\mathbb{Z}^2 \\ |n|\le\varepsilon^{-1}}} g_{1,n} e^{in\cdot x} \Bigg). \tag{1.30}$$

Note that $(\phi^\omega_{\varepsilon,0}, \phi^\omega_{\varepsilon,1})$ converges to $(\phi^\omega_0, \phi^\omega_1)$ in (1.14) distributed according to the Gaussian measure $\mu$ defined in (1.5).

The rescaled model (1.29) on $\mathbb{T}^2$ allows us to study the large temporal and spatial scale behavior of the solution $v_\varepsilon$ to (1.27). Moreover, by suitably choosing $\rho_\varepsilon$, the family $\{u_\varepsilon\}_{\varepsilon>0}$ converges to the solution $u$ to the Wick ordered NLW on $\mathbb{T}^2$ with a parameter $\lambda = \lambda(f)$, depending only on $f$.

**Theorem 1.9.** *Let $\rho > 0$. Then, there exists a choice of $\rho_\varepsilon$ such that, as $\varepsilon \to 0$, the family of the solutions $\{u_\varepsilon\}_{\varepsilon>0}$ to* (1.29) *converges almost surely to the solution $u$ to the following Wick ordered cubic NLW on $\mathbb{T}^2$:*

$$\begin{cases} \partial_t^2 u - \Delta u + \rho u = \lambda :\! u^3 \!: \\ (u, \partial_t u)|_{t=0} = (\phi^\omega_0, \phi^\omega_1), \end{cases} \tag{1.31}$$

*where the convergence takes place in $C([-T_\omega, T_\omega]; H^s(\mathbb{T}^2))$, $s < 0$, for some $T_\omega > 0$. Here, the constant $\lambda$ is given by $\lambda = \frac{f^{(3)}(0)}{6}$, depending only on the function $f$.*

This theorem shows a kind of weak universality for the Wick ordered NLW. See [16] for a similar result for the Wick ordered stochastic NLW. We also refer readers to [19, 17, 18] for more discussion on weak universality (for stochastic parabolic equations, in particular the KPZ equation).

**Remark 1.10.** By starting with the following NLW on $\mathbb{T}^2_\varepsilon$:

$$\begin{cases} \partial_t^2 v_\varepsilon - \Delta v_\varepsilon + \rho_\varepsilon v_\varepsilon + \sum_{j=1}^{m-1} a_j(\varepsilon) v_\varepsilon^{2j+1} = f(v_\varepsilon) \\ (v_\varepsilon, \partial_t v_\varepsilon)|_{t=0} = (\psi^\omega_{\varepsilon,0}, \psi^\omega_{\varepsilon,1}), \end{cases}$$

we can tune the $m$ parameters $\rho_\varepsilon, a_j(\varepsilon)$, $j = 1, \dots, m-1$, such that by a small modification of the proof of Theorem 1.9, we obtain the following Wick ordered NLW:

$$\begin{cases} \partial_t^2 u - \Delta u + \rho u = \lambda :\! u^{2m+1} \!: \\ (u, \partial_t u)|_{t=0} = (\phi^\omega_0, \phi^\omega_1), \end{cases}$$

for some $\lambda = \lambda(f)$, as $\varepsilon \to 0$. In this case, one needs to use the scaling $u_\varepsilon(t,x) = \varepsilon^{-\gamma} v_\varepsilon(\varepsilon^{-1}x, \varepsilon^{-1}t)$ for some suitably chosen $\gamma = \gamma(m) > 0$ and also assume a bound analogous to (1.28) for a higher order derivative of $f$.

## 2. Probabilistic tools

In this section, we first recall basic probabilistic tools. Then, we prove a uniform (in $N$) bound on the Wick ordered monomials $:z_N^\ell: = H_\ell(z_N, \sigma_N)$, consisting of the random linear solution (Proposition 2.3). Moreover, we prove that $\{ :z_N^\ell: \}_{N\in\mathbb{N}}$ is a Cauchy sequence, allowing us to define $:z^\ell:$ by (1.21).

### 2.1. **Hermite polynomials and white noise functional.**

First, recall the Hermite polynomials $H_k(x;\sigma)$ defined via the generating function:

$$F(t,x;\sigma) := e^{tx-\frac{1}{2}\sigma t^2} = \sum_{k=0}^\infty \frac{t^k}{k!} H_k(x;\sigma). \tag{2.1}$$

For simplicity, we set $H_k(x) := H_k(x;1)$. In the following, we list the first few Hermite polynomials for readers' convenience:

$$\begin{aligned} &H_0(x;\sigma) = 1, \qquad H_1(x;\sigma) = x, \qquad H_2(x;\sigma) = x^2 - \sigma, \\ &H_3(x;\sigma) = x^3 - 3\sigma x, \qquad H_4(x;\sigma) = x^4 - 6\sigma x^2 + 3\sigma^2. \end{aligned} \tag{2.2}$$

Next, we define the white noise functional. Let $\xi(x;\omega)$ be the (real-valued) mean-zero Gaussian white noise on $\mathbb{T}^2$ defined by

$$\xi(x;\omega) = \sum_{n\in\mathbb{Z}^2} g_n(\omega) e^{in\cdot x},$$

where $\{g_n\}_{n\in\mathbb{Z}^2}$ is a sequence of independent standard complex-valued Gaussian random variables conditioned that $g_{-n} = \overline{g_n}$, $n\in\mathbb{Z}^2$. It is easy to see that $\xi \in \mathcal{H}^s(\mathbb{T}^2)\setminus\mathcal{H}^{-1}(\mathbb{T}^2)$, $s<-1$, almost surely. In particular, $\xi$ is a distribution, acting on smooth functions. In fact, the action of $\xi$ can be defined on $L^2(\mathbb{T}^2)$.

We define the white noise functional $W_{(\cdot)} : L^2(\mathbb{T}^2) \to L^2(\Omega)$ by

$$W_f(\omega) = \langle f, \xi(\omega)\rangle_{L^2} = \sum_{n\in\mathbb{Z}^2} \widehat{f}(n)\overline{g_n}(\omega) \tag{2.3}$$

for a real-valued function $f\in L^2(\mathbb{T}^2)$. Note that $W_f = \xi(f)$ is basically the Wiener integral of $f$. In particular, $W_f$ is a real-valued Gaussian random variable with mean 0 and variance $\|f\|_{L^2}^2$. Moreover, $W_{(\cdot)}$ is unitary:

$$E\big[W_f W_h\big] = \langle f, h\rangle_{L^2_x} \tag{2.4}$$

for $f,h\in L^2(\mathbb{T}^2)$. The following lemma extends the relation (2.4) to a more general setting.

**Lemma 2.1.** *Let $f,h\in L^2(\mathbb{T}^2)$ such that $\|f\|_{L^2} = \|h\|_{L^2} = 1$. Then, for $k,m\in\mathbb{Z}_{\geq 0}$, we have*

$$\mathbb{E}\big[H_k(W_f)H_m(W_h)\big] = \delta_{km} k! [\langle f,h\rangle]^k.$$

*Here, $\delta_{km}$ denotes the Kronecker's delta function.*

This lemma follows from computing the left-hand side of

$$\mathbb{E}[F(t,W_f)F(s,W_h)] = \sum_{k,m=0}^\infty \frac{t^k}{k!}\frac{s^m}{m!}\mathbb{E}\big[H_k(W_f)H_m(W_h)\big]$$

and comparing the coefficients. See [12, 29] for details.

We also recall the following Wiener chaos estimate [31, Theorem I.22].

**Lemma 2.2.** *Fix $k \in \mathbb{N}$ and $c(n_1, \dots, n_k) \in \mathbb{C}$. Given $d \in \mathbb{N}$, let $\{g_n\}_{n=1}^d$ be a sequence of independent standard complex-valued Gaussian random variables and set $g_{-n} = \overline{g_n}$. Define $S_k(\omega)$ by*

$$S_k(\omega) = \sum_{\Gamma(k,d)} c(n_1, \dots, n_k) g_{n_1}(\omega) \cdots g_{n_k}(\omega),$$

*where $\Gamma(k,d)$ is defined by*

$$\Gamma(k,d) = \big\{(n_1, \dots, n_k) \in \{0, \pm 1, \dots, \pm d\}^k\big\}.$$

*Then, for $p \geq 2$, we have*

$$\|S_k\|_{L^p(\Omega)} \leq (p-1)^{\frac{k}{2}} \|S_k\|_{L^2(\Omega)}. \tag{2.5}$$

The crucial point is that the constant in (2.5) is *independent* of $d \in \mathbb{N}$. This lemma is a direct corollary to the hypercontractivity of the Ornstein-Uhlenbeck semigroup due to Nelson [26].

2.2. **Stochastic estimate on Wick ordered monomials.** In this subsection, we study the Wick ordered monomials $: z_N^\ell :$ and $: z^\ell :$, consisting of the random linear solution $z$ defined in (1.17). From (1.14) and (1.17), we have

$$\widehat{z}(t,n) = \frac{\cos(t\langle n\rangle_\rho)}{\langle n\rangle_\rho} g_{0,n} + \frac{\sin(t\langle n\rangle_\rho)}{\langle n\rangle_\rho} g_{1,n}. \tag{2.6}$$

In order to avoid the combinatorial complexity in higher ordered monomials, we use the white noise functional as in [30]. We, however, need to adapt the white noise functional to $z(t)$. In view of (2.6), we define the white noise functional $W^t_{(\cdot)} : L^2(\mathbb{T}^2) \to L^2(\Omega)$ with a parameter $t \in \mathbb{R}$ by

$$W^t_f(\omega) = \langle f, \xi^t(\omega)\rangle_{L^2} = \sum_{n\in\mathbb{Z}^2} \widehat{f}(n)\overline{g^t_n}(\omega). \tag{2.7}$$

Here, $\xi^t$ denotes (a specific realization of) the white noise on $\mathbb{T}^2$ given by

$$\xi^t(x;\omega) = \sum_{n\in\mathbb{Z}^2} g^t_n(\omega) e^{in\cdot x},$$

where $g^t_n$ is define by

$$g^t_n = \cos(t\langle n\rangle_\rho) g_{0,n} + \sin(t\langle n\rangle_\rho) g_{1,n}.$$

Note that $\mathbb{E}[g^t_n] = 0$ and $\mathrm{Var}(g^t_n) = \cos^2(t\langle n\rangle_\rho) + \sin^2(t\langle n\rangle_\rho) = 1$. Thus, for each fixed $t \in \mathbb{R}$, $\{g^t_n\}_{n\in\mathbb{Z}^2}$ is a sequence of independent standard Gaussian random variables conditioned that $g^t_{-n} = \overline{g^t_n}$ for all $n \in \mathbb{N}$. Therefore, the white noise functional $W^t_{(\cdot)}$ defined in (2.7) satisfies the same properties as the standard white noise functional $W_{(\cdot)}$ defined in (2.3). Lastly, note that, in view of (2.6), the random linear solution $z_N = \mathbf{P}_N z$ can be expressed as

$$z_N(t,x) = \sum_{|n|\leq N} \frac{g^t_n(\omega)}{\langle n\rangle_\rho} e^{in\cdot x}. \tag{2.8}$$

In the following, we use the short-hand notation $L^q_T = L^q_t([-T,T])$, etc.

**Proposition 2.3.** *Let $\ell \in \mathbb{N}$ and $\rho > 0$. Then, given $2 \le q, r < \infty$ and $\varepsilon > 0$, there exist $C, c > 0$ such that*

$$P\Big( \| \langle \nabla \rangle^{-\varepsilon} :z_N^\ell: \|_{L^q_T L^r_x} > \lambda \Big) \le C \exp \bigg( - c \frac{\lambda^{\frac{2}{\ell}}}{T^{\frac{2}{q\ell}}} \bigg) \tag{2.9}$$

*for any $T > 0$, $\lambda > 0$, and any $N \in \mathbb{N}$. Moreover, the sequence $\{ :z_N^\ell: \}_{N \in \mathbb{N}}$ is a Cauchy sequence in $L^p(\Omega; L^q([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2)))$. In particular, denoting the limit by $:z^\ell:$, we have $:z^\ell: \in L^q([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2))$ almost surely, satisfying the tail estimate (2.9).*

Before proceeding to the proof of Proposition 2.3, we introduce some notations. Let $\sigma_N$ be as in (1.8). For *fixed* $x \in \mathbb{T}^2$ and $N \in \mathbb{N}$, we also define

$$\eta_N(x)(\cdot) \stackrel{\text{def}}{=} \frac{1}{\sigma_N^{\frac{1}{2}}} \sum_{|n| \le N} \frac{\overline{e_n(x)}}{\langle n \rangle_\rho} e_n(\cdot) \qquad \text{and} \qquad \gamma_N(\cdot) \stackrel{\text{def}}{=} \sum_{|n| \le N} \frac{1}{\langle n \rangle_\rho^2} e_n(\cdot), \tag{2.10}$$

where $e_n(y) = e^{in \cdot y}$. Note that $\eta_N(x)(\cdot)$ is real-valued with $\|\eta_N(x)\|_{L^2(\mathbb{T}^2)} = 1$ for all $x \in \mathbb{T}^2$ and all $N \in \mathbb{N}$. Moreover, we have

$$\langle \eta_M(x), \eta_N(y) \rangle_{L^2} = \frac{1}{\sigma_M^{\frac{1}{2}} \sigma_N^{\frac{1}{2}}} \gamma_N(y - x) = \frac{1}{\sigma_M^{\frac{1}{2}} \sigma_N^{\frac{1}{2}}} \gamma_N(x - y), \tag{2.11}$$

for fixed $x, y \in \mathbb{T}^2$ and $M \ge N \ge 1$.

*Proof.* From (2.8) and (2.10), we have

$$z_N(t,x) = \sigma_N^{\frac{1}{2}} \frac{z_N(t,x)}{\sigma_N^{\frac{1}{2}}} = \sigma_N^{\frac{1}{2}} W^t_{\overline{\eta_N(x)}} = \sigma_N^{\frac{1}{2}} W^t_{\eta_N(x)}. \tag{2.12}$$

Then, from (1.18) and (2.12), we have

$$:z_N^\ell(t,x): = H_\ell(z_N(t,x); \sigma_N) = \sigma_N^{\frac{\ell}{2}} H_\ell\big(W^t_{\eta_N(x)}\big). \tag{2.13}$$

Given $n \in \mathbb{Z}^2$, define $\Gamma_\ell(n)$ by

$$\Gamma_\ell(n) = \{ (n_1, \dots, n_\ell) \in (\mathbb{Z}^2)^\ell : n_1 + \cdots + n_\ell = n \}.$$

Then, for $(n_1, \dots, n_\ell) \in \Gamma_\ell(n)$, we have $\max_j |n_j| \gtrsim |n|$. Thus, it follows from Lemma 2.1 with (2.13) and (2.11) that

$$\begin{aligned}
\| \langle :z_N^\ell(t):, e_n \rangle \|^2_{L^2(\Omega)} &= \sigma_N^\ell \int_{\mathbb{T}^2_x \times \mathbb{T}^2_y} \overline{e_n(x)} e_n(y) \int_\Omega H_\ell\big(W^t_{\eta_N(x)}\big) \overline{H_\ell\big(W^t_{\eta_N(y)}\big)} dP dx dy \\
&= \ell! \int_{\mathbb{T}^2_x \times \mathbb{T}^2_y} [\gamma_N(x-y)]^\ell \overline{e_n(x-y)} dx dy \\
&= \ell! \cdot \mathcal{F}[\gamma_N^\ell](n) = \ell! \sum_{\substack{\Gamma_\ell(n) \\ |n_j| \le N}} \prod_{j=1}^\ell \frac{1}{\langle n_j \rangle_\rho^2} \lesssim \frac{1}{\langle n \rangle^{2(1-\theta)}}
\end{aligned} \tag{2.14}$$

for any $\theta > 0$. On the other hand, for $n \neq n'$, we have

$$\begin{aligned}
&\int_\Omega \langle :z_N^\ell(t):, e_n\rangle \overline{\langle :z_N^\ell(t):, e_{n'}\rangle} dP \\
&\qquad = \sigma_N^\ell \int_{\mathbb{T}^2_x\times\mathbb{T}^2_y} \overline{e_n(x)} e_{n'}(y) \int_\Omega H_\ell\big(W^t_{\eta_N(x)}\big)\overline{H_\ell\big(W^t_{\eta_N(y)}\big)} dP dx dy \\
&\qquad = \ell! \int_{\mathbb{T}^2_x}\int_{\mathbb{T}^2_y} [\gamma_N(x-y)]^\ell \overline{e_{n'}(x-y)} dy\, \overline{e_n(x)} e_{n'}(x) dx \\
&\qquad = \ell! \cdot \mathcal{F}[\gamma_N^\ell](n') \int_{\mathbb{T}^2_x} \overline{e_n(x)} e_{n'}(x) dx = 0. \qquad (2.15)
\end{aligned}$$

Hence, given $x \in \mathbb{T}^2$ and $t \in \mathbb{R}$, it follows from (2.14) and (2.15) that

$$\begin{aligned}
\| \langle \nabla\rangle^{-\varepsilon} :z_N^\ell(t,x): \|_{L^2(\Omega)} &= \bigg\| \sum_{n\in\mathbb{Z}^2} \langle n\rangle^{-\varepsilon} \mathcal{F}_x\big[\ :z_N^\ell(t):\big](n)\, e^{in\cdot x} \bigg\|_{L^2(\Omega)} \\
&\le C_\ell \bigg( \sum_{n\in\mathbb{Z}^2} \langle n\rangle^{-2\varepsilon} \mathcal{F}[\gamma_N^\ell](n)\bigg)^{\frac12} \le C_\ell \bigg( \sum_{n\in\mathbb{Z}^2} \langle n\rangle^{-2(1+\varepsilon-\theta)}\bigg)^{\frac12} \\
&< \infty, \qquad (2.16)
\end{aligned}$$

uniformly in $N \in \mathbb{N}$, as long as $0 < \theta < \varepsilon$.

Fix $2 \le q, r < \infty$. Then, by Minkowski's integral inequality, Lemma 2.2 (with (2.8)), and (2.16), we have

$$\begin{aligned}
\Big\| \| \langle \nabla\rangle^{-\varepsilon} :z_N^\ell: \|_{L^q_T L^r_x} \Big\|_{L^p(\Omega)} &\le \Big\| \| \langle \nabla\rangle^{-\varepsilon} :z_N^\ell(t,x): \|_{L^p(\Omega)} \Big\|_{L^q_T L^r_x} \\
&\le C_\ell\, p^{\frac{\ell}{2}} \Big\| \| \langle \nabla\rangle^{-\varepsilon} :z_N^\ell(t,x): \|_{L^2(\Omega)} \Big\|_{L^q_T L^r_x} \\
&\lesssim T^{\frac1q} p^{\frac{\ell}{2}}, \qquad (2.17)
\end{aligned}$$

for all $p \ge \max(q, r)$. Finally, (2.9) follows from (2.17) and Chebyshev's inequality.

A similar computation with Lemma 2.1, (2.11), and Lemma 2.2 shows that the sequence $\big\{ :z_N^\ell: \big\}_{N\in\mathbb{N}}$ is a Cauchy sequence in $L^p(\Omega; L^q([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2)))$. □

**Remark 2.4.** As a corollary to Proposition 2.3, we can show that the tail estimate (2.9) and the convergence of $:z_N^\ell:$ to $:z^\ell:$ hold even when $q = \infty$ and/or $r = \infty$. This follows from applying Sobolev's inequality (in time and/or space) and using the fact that $z$ solves the linear wave/Klein-Gordon equation. See [7]. With this observation, we can easily show that $:z_N^\ell:,\ :z^\ell: \in C([-T,T]; W^{-\varepsilon,r}(\mathbb{T}^2))$ almost surely. See also [16, 27].

## 3. Local well-posedness of the Wick ordered NLW

In this section, we present the proof of Theorem 1.2. We combine the deterministic analysis via the Fourier restriction norm method (with the hyperbolic Sobolev spaces) and the stochastic estimate on the Wick ordered monomials $:z^\ell:$ (Proposition 2.3). In the following, we fix $\rho > 0$.

3.1. **Hyperbolic Sobolev spaces and Strichartz estimates.** We first recall the hyperbolic Sobolev space $X^{s,b}$ due to Klainerman-Machedon [22] and Bourgain [2], defined by the norm

$$\|u\|_{X^{s,b}(\mathbb{R}\times\mathbb{T}^2)} = \|\langle n\rangle^s \langle |\tau| - \langle n\rangle_\rho\rangle^b \widehat{u}(\tau,n)\|_{\ell^2_n L^2_\tau(\mathbb{R}\times\mathbb{Z}^2)}.$$

For $b > \frac{1}{2}$, we have $X^{s,b} \subset C(\mathbb{R}; H^s)$. Given an interval $I \subset \mathbb{R}$, we define the local-in-time version $X^{s,b}(I)$ as a restriction norm:

$$\|u\|_{X^{s,b}(I)} = \inf\big\{\|v\|_{X^{s,b}(\mathbb{R}\times\mathbb{T}^2)} : v|_I = u\big\}.$$

When $I = [-T,T]$, we set $X^{s,b}_T = X^{s,b}(I)$.

The main deterministic tool for the proof of Theorem 1.2 is the following Strichartz estimates for the linear wave/Klein-Gordon equation. Given $0 \le s \le 1$, we say that a pair $(q,r)$ is $s$-admissible if $2 < q \le \infty$, $2 \le r < \infty$,

$$\frac{1}{q} + \frac{2}{r} = 1 - s, \qquad \text{and} \qquad \frac{1}{q} + \frac{1}{2r} \le \frac{1}{4}.$$

Then, we have the following Strichartz estimates.

**Lemma 3.1.** *Let $T \le 1$. Given $0 \le s \le 1$, let $(q,r)$ be $s$-admissible. Then, we have*

$$\|S(t)(\phi_0,\phi_1)\|_{L^q_T L^r_x(\mathbb{T}^2)} \lesssim \|(\phi_0,\phi_1)\|_{\mathcal{H}^s(\mathbb{T}^2)}. \tag{3.1}$$

See Ginibre-Velo [14], Lindblad-Sogge [24], and Keel-Tao [20] for the Strichartz estimates on $\mathbb{R}^d$. See also [21]. The Strichartz estimates (3.1) on $\mathbb{T}^2$ in Lemma 3.1 follows from those on $\mathbb{R}^2$ and the finite speed of propagation.

When $b > \frac{1}{2}$, the $X^{s,b}$-spaces enjoy the transference principle. In particular, as a corollary to Lemma 3.1, we obtain the following space-time estimate. See [23, 32] for the proof.

**Lemma 3.2.** *Let $T \le 1$. Given $0 \le s \le 1$, let $(q,r)$ be $s$-admissible. Then, for $b > \frac{1}{2}$, we have*

$$\|u\|_{L^q_T L^r_x} \lesssim \|u\|_{X^{s,b}_T}.$$

Lastly, we state the nonhomogeneous linear estimate. See [13].

**Lemma 3.3.** *Let $-\frac{1}{2} < b' \le 0 \le b \le b' + 1$. Then, for $T \le 1$, we have*

$$\bigg\| \int_0^t \frac{\sin((t-t')\langle\nabla\rangle_\rho)}{\langle\nabla\rangle_\rho} F(t')dt' \bigg\|_{X^{s,b}_T} \lesssim T^{1-b+b'} \|F\|_{X^{s-1,b'}_T}.$$

3.2. **Proof of Theorem 1.2.** In the following, we simply consider the case $s = s_{\text{crit}} + \delta$ with $\delta \ll 1$. Given $T \le 1$, define $\Psi(w)$ by

$$\Psi(w)(t) = \Psi^\omega(w)(t) \overset{\text{def}}{=} \int_0^t \frac{\sin((t-t')\langle\nabla\rangle_\rho)}{\langle\nabla\rangle_\rho} :(w+z)^{2m+1}(t'):\, dt'.$$

Let $b = \frac{1}{2}+$. Then, for $0 < \theta \le 1 - b$, by Lemma 3.3, we have

$$\|\Psi(w)\|_{X^{s,b}_T} \lesssim T^\theta \| :(w+z)^{2m+1}: \|_{X^{s-1,b-1+\theta}_T}. \tag{3.2}$$

From (1.23), we have

$$:(w+z)^{2m+1}: = \sum_{\ell=0}^{2m+1} \binom{2m+1}{\ell} w^{2m+1-\ell} :z^\ell: .$$

Then, by duality, we have

$$\begin{aligned} \| :(w+z)^{2m+1}: \|_{X_T^{s-1,b-1+\theta}} &\le \sum_{\ell=0}^{2m+1} C_{m,\ell} \| w^{2m+1-\ell} :z^\ell: \|_{X_T^{s-1,b-1+\theta}} \\ &\le \sum_{\ell=0}^{2m+1} C_{m,\ell} \sup_{h_\ell} \bigg| \iint \mathbf{1}_{[-T,T]} \widetilde{w}^{2m+1-\ell} :z^\ell: h_\ell \, dxdt \bigg|, \end{aligned} \tag{3.3}$$

for any extension $\widetilde{w}$ of $w$, where the supremum is taken over $h_\ell \in X^{1-s,1-b-\theta}$ with $\|h_\ell\|_{X^{1-s,1-b-\theta}} = 1$. By choosing $\theta > 0$ sufficiently small, we have $1-b-\theta = \frac{1}{2}-$.

• **Case 1:** $m = 1$.

In this case, we have $s = s_{\text{crit}} + \delta = \frac{1}{4} + \delta$. Noting that $\big(\frac{12}{1+2\delta}, \frac{3}{1-\delta}\big)$ is $\big(\frac{1}{4} + \frac{1}{2}\delta\big)$-admissible, it follows from Lemma 3.2 that

$$\|\langle \nabla \rangle^\varepsilon \widetilde{w}\|_{L_T^{\frac{12}{1+2\delta}} L_x^{\frac{3}{1-\delta}}} \lesssim \|\widetilde{w}\|_{X^{\frac{1}{4}+\frac{1}{2}\delta+\varepsilon, \frac{1}{2}+}} \lesssim \|\widetilde{w}\|_{X^{s,\frac{1}{2}+}}$$

for any extension $\widetilde{w}$ of $w$, as long as $\varepsilon \le \frac{1}{2}\delta$. By taking an infimum over all the extensions $\widetilde{w}$ of $w$, we obtain

$$\inf_{\widetilde{w}|_{[-T,T]} = w} \|\langle \nabla \rangle^\varepsilon \widetilde{w}\|_{L_T^{\frac{12}{1+2\delta}} L_x^{\frac{3}{1-\delta}}} \lesssim \|w\|_{X_T^{s,\frac{1}{2}+}}. \tag{3.4}$$

On the one hand, noting that $\big(\frac{4}{1-2\delta}, \frac{1}{\delta}\big)$ is $\big(\frac{3}{4} - \frac{3}{2}\delta\big)$-admissible, Hölder's inequality (with $T \le 1$) and Lemma 3.2 yield

$$\|\langle \nabla \rangle^\varepsilon h_\ell\|_{L_T^{\frac{4}{3-2\delta}} L_x^{\frac{1}{\delta}}} \lesssim \|\langle \nabla \rangle^\varepsilon h_\ell\|_{L_T^{\frac{4}{1-2\delta}} L_x^{\frac{1}{\delta}}} \lesssim \|\langle \nabla \rangle^\varepsilon h_\ell\|_{X^{\frac{3}{4}-\frac{3}{2}\delta, \frac{1}{2}+}}. \tag{3.5}$$

On the other hand, applying Hölder's inequality in $t$ and Sobolev's inequality in $x$, we have

$$\|\langle \nabla \rangle^\varepsilon h_\ell\|_{L_T^{\frac{4}{3-2\delta}} L_x^{\frac{1}{\delta}}} \lesssim \|\langle \nabla \rangle^\varepsilon h_\ell\|_{L_T^2 L_x^{\frac{1}{\delta}}} \lesssim \|\langle \nabla \rangle^\varepsilon h_\ell\|_{X^{1-2\delta,0}}. \tag{3.6}$$

Interpolating (3.5) and (3.6) with sufficiently small $\theta > 0$, we obtain

$$\|\langle \nabla \rangle^\varepsilon h_\ell\|_{L_T^{\frac{4}{3-2\delta}} L_x^{\frac{1}{\delta}}} \lesssim \|h_\ell\|_{X^{\frac{3}{4}-\frac{5}{4}\delta+\varepsilon, 1-b-\theta}} \lesssim \|h_\ell\|_{X^{1-s,1-b-\theta}} \tag{3.7}$$

as long as $\varepsilon \le \frac{1}{4}\delta$.

For $\ell = 0, 1, 2, 3$, define $(q_\ell, r_\ell)$ by

$$1 = (3-\ell)\frac{1+2\delta}{12} + \frac{3-2\delta}{4} + \frac{1}{q_\ell} \qquad \text{and} \qquad 1 = (3-\ell)\frac{1-\delta}{3} + \delta + \frac{1}{r_\ell}.$$

When $\ell = 0$, we have $q_0 = r_0 = \infty$ and $:z^0: \equiv 1$. Then, by fractional Leibniz rule and Hölder's inequality with (3.4) and (3.7), we have

$$\begin{aligned}
\inf_{\widetilde{w}|_{[-T,T]}=w} \bigg| \iint \mathbf{1}_{[-T,T]} \widetilde{w}^{3-\ell} :z^\ell: h_\ell \, dxdt \bigg| \\
&= \inf_{\widetilde{w}|_{[-T,T]}=w} \bigg| \iint \mathbf{1}_{[-T,T]} \langle \nabla \rangle^{\varepsilon} (\widetilde{w}^{3-\ell} h_\ell) \langle \nabla \rangle^{-\varepsilon} :z^\ell: \, dxdt \bigg| \\
&\lesssim \inf_{\widetilde{w}|_{[-T,T]}=w} \| \langle \nabla \rangle^{\varepsilon} \widetilde{w} \|^{3-\ell}_{L_T^{\frac{12}{1+2\delta}} L_x^{\frac{3}{1-\delta}}} \| \langle \nabla \rangle^{\varepsilon} h_\ell \|_{L_T^{\frac{4}{3-2\delta}} L_x^{\frac{1}{\delta}}} \| \langle \nabla \rangle^{-\varepsilon} :z^\ell: \|_{L_T^{q_\ell} L_x^{r_\ell}} \\
&\lesssim \|w\|^{3-\ell}_{X_T^{s,b}} \|h_\ell\|_{X^{1-s,1-b-\theta}} \| \langle \nabla \rangle^{-\varepsilon} :z^\ell: \|_{L_T^{q_\ell} L_x^{r_\ell}} \\
&= \|w\|^{3-\ell}_{X_T^{s,b}} \| \langle \nabla \rangle^{-\varepsilon} :z^\ell: \|_{L_T^{q_\ell} L_x^{r_\ell}}
\end{aligned} \tag{3.8}$$

as long as $0 < \varepsilon \le \frac{1}{4}\delta$. Hence, by Proposition 2.3 with (3.2), (3.3), and (3.8), we obtain

$$\|\Psi(w)\|_{X_T^{s,b}} \lesssim T^\theta \sum_{\ell=0}^{3} \|w\|^{3-\ell}_{X_T^{s,b}} \tag{3.9}$$

outside a set of probability $< \exp\big(-\frac{1}{T^c}\big)$ for some $c > 0$. Similarly, we have

$$\|\Psi(w_1) - \Psi(w_2)\|_{X_T^{s,b}} \lesssim T^\theta \sum_{\ell=0}^{2} \big( \|w_1\|^{2-\ell}_{X_T^{s,b}} + \|w_2\|^{2-\ell}_{X_T^{s,b}} \big) \|w_1 - w_2\|_{X_T^{s,b}} \tag{3.10}$$

outside a set of probability $< \exp\big(-\frac{1}{T^c}\big)$. Therefore, it follows from (3.9) and (3.10) that for each $T \ll 1$, there exists a set $\Omega_T$ with $P(\Omega_T^c) < \exp\big(-\frac{1}{T^c}\big)$ such that, for each $\omega \in \Omega_T$, $\Psi^\omega$ is a contraction on a ball of radius $O(1)$ in $X_T^{s,b}$.

- **Case 2:** $m \ge 2$.

In this case, we have $s = s_{\text{crit}} + \delta = 1 - \frac{1}{m} + \delta$. Define $(q, r)$ by

$$\frac{1}{q} = \frac{3m-1}{3m(2m+1)} + \frac{\delta}{6} \qquad \text{and} \qquad \frac{1}{r} = \frac{3m+4}{6m(2m+1)} - \frac{\delta}{3}.$$

Noting that $(q, r)$ is $\big(s_{\text{crit}} + \frac{1}{2}\delta\big)$-admissible, it follows from Lemma 3.2 that

$$\| \langle \nabla \rangle^{\varepsilon} \widetilde{w} \|_{L_T^q L_x^r} \lesssim \|\widetilde{w}\|_{X^{s_{\text{crit}} + \frac{1}{2}\delta + \varepsilon, \frac{1}{2}+}} \lesssim \|\widetilde{w}\|_{X^{s, \frac{1}{2}+}}$$

for any extension $\widetilde{w}$ of $w$, as long as $\varepsilon \le \frac{1}{2}\delta$. By taking an infimum over all the extensions $\widetilde{w}$ of $w$, we obtain

$$\inf_{\widetilde{w}|_{[-T,T]}=w} \| \langle \nabla \rangle^{\varepsilon} \widetilde{w} \|_{L_T^q L_x^r} \lesssim \|w\|_{X_T^{s, \frac{1}{2}+}}. \tag{3.11}$$

Now, define $(\widetilde{q}, \widetilde{r})$ by

$$\frac{1}{\widetilde{q}} = \frac{1}{3m} - \frac{2m+1}{6}\delta \qquad \text{and} \qquad \frac{1}{\widetilde{r}} = \frac{3m-4}{6m} + \frac{2m+1}{3}\delta.$$

Then, $(\widetilde{q}, \widetilde{r})$ is $\big(1 - s_{\text{crit}} - \frac{2m+1}{2}\delta\big)$-admissible. On the one hand, by Lemma 3.2, we have

$$\| \langle \nabla \rangle^{\varepsilon} h_\ell \|_{L_T^{\widetilde{q}} L_x^{\widetilde{r}}} \lesssim \| \langle \nabla \rangle^{\varepsilon} h_\ell \|_{X^{1 - s_{\text{crit}} - \frac{2m+1}{2}\delta, \frac{1}{2}+}}. \tag{3.12}$$

On the other hand, by Sobolev's inequality, we have

$$\| \langle \nabla \rangle^{\varepsilon} h_\ell \|_{L_T^{\widetilde{q}} L_x^{\widetilde{r}}} \lesssim \| \langle \nabla \rangle^{\varepsilon} h_\ell \|_{X^{1 - \frac{3m-4}{3m} - \frac{4m+2}{3}\delta, \frac{1}{2} - \frac{1}{3m} + \frac{2m+1}{6}\delta}}. \tag{3.13}$$

Note that the temporal regularity on the right-hand side of (3.13) is less than $\frac{1}{2}$ by choosing $\delta > 0$ sufficiently small. Hence, by interpolating (3.12) and (3.13) with sufficiently small $\theta > 0$, we obtain

$$\|\langle\nabla\rangle^{\varepsilon} h_\ell\|_{L^{\widetilde{q}}_T L^{\widetilde{r}}_x} \lesssim \|h_\ell\|_{X^{1-s_{\mathrm{crit}}-m\delta+\varepsilon,1-b-\theta}} \lesssim \|h_\ell\|_{X^{1-s,1-b-\theta}} \tag{3.14}$$

as long as $\varepsilon \le (m-1)\delta$.

Proceeding as before, it follows from Hölder's inequality with (3.11) and (3.14) that

$$\begin{aligned}
&\bigg| \iint \inf_{\widetilde{w}|_{[-T,T]}=w} \mathbf{1}_{[-T,T]} \widetilde{w}^{2m+1-\ell} :z^\ell: h_\ell \, dxdt \bigg| \\
&\quad = \inf_{\widetilde{w}|_{[-T,T]}=w} \bigg| \iint \mathbf{1}_{[-T,T]} \langle\nabla\rangle^{\varepsilon}(\widetilde{w}^{2m+1-\ell} h_\ell) \langle\nabla\rangle^{-\varepsilon} :z^\ell: \, dxdt \bigg| \\
&\quad \lesssim \inf_{\widetilde{w}|_{[-T,T]}=w} \|\langle\nabla\rangle^{\varepsilon}\widetilde{w}\|^{2m+1-\ell}_{L^q_T L^r_x} \|\langle\nabla\rangle^{\varepsilon} h_\ell\|_{L^{\widetilde{q}}_T L^{\widetilde{r}}_x} \|\langle\nabla\rangle^{-\varepsilon} :z^\ell:\|_{L^{\frac{q}{\ell}}_T L^{\frac{r}{\ell}}_x} \\
&\quad \lesssim \|w\|^{2m+1-\ell}_{X^{s,b}_T} \|h_\ell\|_{X^{1-s,1-b-\theta}} \|\langle\nabla\rangle^{-\varepsilon} :z^\ell:\|_{L^{\frac{q}{\ell}}_T L^{\frac{r}{\ell}}_x} \\
&\quad = \|w\|^{2m+1-\ell}_{X^{s,b}_T} \|\langle\nabla\rangle^{-\varepsilon} :z^\ell:\|_{L^{\frac{q}{\ell}}_T L^{\frac{r}{\ell}}_x}
\end{aligned} \tag{3.15}$$

as long as $0 < \varepsilon \le \frac{1}{2}\delta$. Hence, by Proposition 2.3 with (3.2), (3.3), and (3.15), we obtain

$$\begin{aligned}
\|\Psi(w)\|_{X^{s,b}_T} &\lesssim T^\theta \sum_{\ell=0}^{2m+1} \|w\|^{2m+1-\ell}_{X^{s,b}_T}, \\
\|\Psi(w_1) - \Psi(w_2)\|_{X^{s,b}_T} &\lesssim T^\theta \sum_{\ell=0}^{2m} \big(\|w_1\|^{2m-\ell}_{X^{s,b}_T} + \|w_2\|^{2m-\ell}_{X^{s,b}_T}\big) \|w_1 - w_2\|_{X^{s,b}_T}
\end{aligned}$$

outside a set of probability $< \exp\big(-\frac{1}{T^c}\big)$ for some $c > 0$. Therefore, for each $T \ll 1$, there exists a set $\Omega_T$ with $P(\Omega_T^c) < \exp\big(-\frac{1}{T^c}\big)$ such that, for each $\omega \in \Omega_T$, $\Psi^\omega$ is a contraction on a ball of radius $O(1)$ in $X^{s,b}_T$.

This completes the proof of Theorem 1.2.

## 4. Weak universality: Wick ordered NLW as a scaling limit

In this section, we present the proof of Theorem 1.9. We follow closely the argument in [16]. With $z_\varepsilon = z_\varepsilon^\omega = S(t)(\phi^\omega_{\varepsilon,0}, \phi^\omega_{\varepsilon,1})$, let us decompose $u_\varepsilon = z_\varepsilon + w_\varepsilon$ as in (1.22). Then, the residual term $w_\varepsilon$ satisfies

$$\partial_t^2 w_\varepsilon - \Delta w_\varepsilon + \rho w_\varepsilon = F_\varepsilon(w_\varepsilon), \tag{4.1}$$

where $F_\varepsilon(w_\varepsilon)$ is given by

$$\begin{aligned}
F_\varepsilon(w_\varepsilon) &= \varepsilon^{-3}\big\{ f(\varepsilon(z_\varepsilon + w_\varepsilon)) + \varepsilon(\varepsilon^2\rho - \rho_\varepsilon)(z_\varepsilon + w_\varepsilon) \big\} \\
&= \varepsilon^{-2}\{f'(0) + \varepsilon^2\rho - \rho_\varepsilon\}(z_\varepsilon + w_\varepsilon) + \frac{f^{(3)}(0)}{6}(z_\varepsilon + w_\varepsilon)^3 + R_\varepsilon,
\end{aligned} \tag{4.2}$$

where the second equality follows from $f(0) = f''(0) = 0$ and Taylor's remainder theorem with the remainder term $R_\varepsilon$ given by

$$R_\varepsilon = \varepsilon \int_0^1 \frac{(1-\theta)^3}{6} f^{(4)}(\theta\varepsilon(z_\varepsilon + w_\varepsilon)) d\theta \cdot (z_\varepsilon + w_\varepsilon)^4. \tag{4.3}$$

From (1.30), we see that $z_\varepsilon(t,x)$ is a mean-zero real-valued Gaussian random variable with variance

$$\sigma_\varepsilon = \mathbb{E}[z_\varepsilon^2(t,x)] \sim \log \varepsilon^{-1}.$$

Note that $\sigma_\varepsilon$ is independent of $x \in \mathbb{T}^2$ and $t \in \mathbb{R}$. Recalling from (2.2) that $x^3 = H_3(x;\sigma) + 3\sigma x$, it follows from (4.2) and (4.3) that

$$\begin{aligned} F_\varepsilon(w_\varepsilon) = \varepsilon^{-2}\bigg\{ f'(0) + \varepsilon^2\rho - \rho_\varepsilon + 3\varepsilon^2\sigma_\varepsilon \frac{f^{(3)}(0)}{6}\bigg\}(z_\varepsilon + w_\varepsilon) \\ + \frac{f^{(3)}(0)}{6} H_3(z_\varepsilon + w_\varepsilon; \sigma_\varepsilon) + R_\varepsilon. \end{aligned}$$

For each $\varepsilon > 0$, we set $\rho_\varepsilon$ by

$$\rho_\varepsilon = f'(0) + \varepsilon^2\rho + \varepsilon^2\sigma_\varepsilon \frac{f^{(3)}(0)}{2}$$

so that the first term on the right-hand side vanishes. Then, by letting $\lambda = \frac{f^{(3)}(0)}{6}$, we obtain

$$F_\varepsilon(w_\varepsilon) = \lambda H_3(z_\varepsilon + w_\varepsilon; \sigma_\varepsilon) + R_\varepsilon \stackrel{\text{def}}{=} \lambda :u_\varepsilon^3: + R_\varepsilon.$$

From (4.3) and (1.28), we have

$$\begin{aligned} |R_\varepsilon| &= \bigg| \varepsilon \int_0^1 \frac{(1-\theta)^3}{6} f^{(4)}(\theta\varepsilon(z_\varepsilon + w_\varepsilon)) d\theta \cdot (z_\varepsilon + w_\varepsilon)^4 \bigg| \\ &\lesssim \varepsilon \big\{ |z_\varepsilon| + |w_\varepsilon| \big\}^{M+4}. \end{aligned}$$

In particular, we can write (4.1) as

$$\partial_t^2 w_\varepsilon - \Delta w_\varepsilon + \rho w_\varepsilon = \lambda \sum_{\ell=0}^{3} \binom{3}{\ell} :z_\varepsilon^\ell: w_\varepsilon^{3-\ell} + O\big(\varepsilon\big\{|z_\varepsilon| + |w_\varepsilon|\big\}^{M+4}\big). \tag{4.4}$$

It follows from Proposition 2.3 with (1.30) that

$$\varepsilon \|z_\varepsilon\|_{L_T^q L_x^r}^{M+4} = o_\varepsilon(1)$$

almost surely. Then, by proceeding as in Section 3 (where we handle the second term on the right-hand side of (4.4) by applying the argument in Section 3 with $2m+1 \geq M+4$), we obtain an a priori bound on $w_\varepsilon$, uniformly in $\varepsilon > 0$. Moreover, the local existence time $T = T_\omega$ can be chosen to be independent of $\varepsilon > 0$.

Let $u$ be the solution to (1.31). In an analogous manner, we can estimate the difference $w - w_\varepsilon$, where $w = u - z$ as in (1.22). Together with the almost sure convergence of $z_\varepsilon$ to $z$ (see Remark 2.4), we see that $u_\varepsilon$ converges to $u$ in $C([-T_\omega, T_\omega]; H^s(\mathbb{T}^2))$ for $s < 0$. This completes the proof of Theorem 1.9.

**Acknowledgements.** T.O. is supported by the ERC starting grant no. 637995 "ProbDyn-DispEq". L.T. is supported by the grants "BEKAM" ANR-15-CE40-0001 and "ISDEEC" ANR-16-CE40-0013.

## References


[1] S. Albeverio, A. Cruzeiro, *Global flows with invariant (Gibbs) measures for Euler and Navier-Stokes two dimensional fluids,* Comm. Math. Phys. 129 (1990) 431–444.

[2] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, I: Schrödinger equations,* Geom. Funct. Anal. 3 (1993), 107–156.

[3] J. Bourgain, *Periodic nonlinear Schrödinger equation and invariant measures*, Comm. Math. Phys. 166 (1994), no. 1, 1–26.

[4] J. Bourgain, *Invariant measures for the 2D-defocusing nonlinear Schrödinger equation*, Comm. Math. Phys. 176 (1996), no. 2, 421–445.

[5] J. Bourgain, *Nonlinear Schrödinger equations,* Hyperbolic equations and frequency interactions (Park City, UT, 1995), 3–157, IAS/Park City Math. Ser., 5, Amer. Math. Soc., Providence, RI, 1999.

[6] D. Brydges, G. Slade, *Statistical mechanics of the 2-dimensional focusing nonlinear Schrödinger equation,* Comm. Math. Phys. 182 (1996), no. 2, 485–504.

[7] N. Burq, L. Thomann, N. Tzvetkov, *Global infinite energy solutions for the cubic wave equation,* Bull. Soc. Math. France 143 (2015), no. 2, 301–313.

[8] N. Burq, L. Thomann, N. Tzvetkov, *Remarks on the Gibbs measures for nonlinear dispersive equations*, Ann. Fac. Sci. Toulouse Math., to appear.

[9] N. Burq, N. Tzvetkov, *Random data Cauchy theory for supercritical wave equations. I. Local theory,* Invent. Math. 173 (2008), no. 3, 449–475.

[10] N. Burq, N. Tzvetkov, *Random data Cauchy theory for supercritical wave equations. II. A global existence result,* Invent. Math. 173 (2008), no. 3, 477–496.

[11] G. Da Prato, A. Debussche, *Two-dimensional Navier-Stokes equations driven by a space-time white noise*, J. Funct. Anal. 196 (2002), no. 1, 180–210.

[12] G. Da Prato, L. Tubaro, *Wick powers in stochastic PDEs: an introduction,* Technical Report UTM, 2006, 39 pp.

[13] J. Ginibre, Y. Tsutsumi, G. Velo, *On the Cauchy problem for the Zakharov system,* J. Funct. Anal. 151 (1997), no. 2, 384–436.

[14] J. Ginibre, G. Velo, *Generalized Strichartz inequalities for the wave equation,* J. Funct. Anal. 133 (1995), 50–68.

[15] J. Glimm, A. Jaffe, *Quantum physics. A functional integral point of view,* Second edition. Springer-Verlag, New York, 1987. xxii+535 pp.

[16] M. Gubinelli, H. Koch, T. Oh, *Renormalization of the two-dimensional stochastic nonlinear wave equations*, arXiv:1703.05461 [math.AP].

[17] M. Gubinelli, N. Perkowski, *KPZ reloaded,*, Commun. Math. Phys. 349 (2017), 165–269.

[18] M. Gubinelli, N. Perkowski, *The Hairer–Quastel universality result at stationarity*, to appear in RIMS Kôkyûroku Bessatsu (2016).

[19] M. Hairer, J. Quastel, *A class of growth models rescaling to KPZ,* arXiv:1512.07845 [math-ph].

[20] M. Keel, T. Tao, *Endpoint Strichartz estimates*, Amer. J. Math. 120 (1998), no. 5, 955–980.

[21] R. Killip, B. Stovall, M. Vişan, *Blowup behaviour for the nonlinear Klein-Gordon equation*, Math. Ann. 358 (2014), no. 1-2, 289–350.

[22] S. Klainerman, M. Machedon, Space-time estimates for null forms and the local existence theorem, Comm. Pure Appl. Math., 46 (1993), 1221–1268.

[23] S. Klainerman, S. Selberg, *Bilinear estimates and applications to nonlinear wave equations,* Commun. Contemp. Math. 4 (2002), no. 2, 223–295.

[24] H. Lindblad, C. Sogge, *On existence and scattering with minimal regularity for semilinear wave equations,* J. Funct. Anal. 130 (1995), 357–426.

[25] H.P. McKean, *Statistical mechanics of nonlinear wave equations. IV. Cubic Schrödinger,* Comm. Math. Phys. 168 (1995), no. 3, 479–491. *Erratum: Statistical mechanics of nonlinear wave equations. IV. Cubic Schrödinger*, Comm. Math. Phys. 173 (1995), no. 3, 675.

[26] E. Nelson, *A quartic interaction in two dimensions*, 1966 Mathematical Theory of Elementary Particles (Proc. Conf., Dedham, Mass., 1965) pp. 69–73 M.I.T. Press, Cambridge, Mass.

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

[28] T. Oh, J. Quastel, *On Cameron-Martin theorem and almost sure global existence*, Proc. Edinb. Math. Soc. 59 (2016), 483–501.

[29] T. Oh, G. Richards, L. Thomann, *On invariant Gibbs measures for the generalized KdV equations*, Dyn. Partial Differ. Equ. 13 (2016), no.2, 133–153.

[30] T. Oh, L. Thomann, *A pedestrian approach to the invariant Gibbs measure for the 2-d defocusing nonlinear Schrödinger equations*, arXiv:1509.02093 [math.AP].

[31] B. Simon, *The* $P(\varphi)_2$ *Euclidean (quantum) field theory,* Princeton Series in Physics. Princeton University Press, Princeton, N.J., 1974. xx+392 pp.

[32] T. Tao, *Nonlinear dispersive equations. Local and global analysis,* CBMS Regional Conference Series in Mathematics, 106. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. xvi+373 pp.

[33] M. Zworski, *Semiclassical analysis.* Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012. xii+431 pp.

Tadahiro Oh, School of Mathematics, The University of Edinburgh, and The Maxwell Institute for the Mathematical Sciences, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, United Kingdom

*E-mail address*: hiro.oh@ed.ac.uk

Laurent Thomann, Institut Élie Cartan, Université de Lorraine, B.P. 70239, F-54506 Vandœuvre-lès-Nancy Cedex, France

*E-mail address*: laurent.thomann@univ-lorraine.fr