J. Eur. Math. Soc. 26, 817–874 (2024) DOI 10.4171/JEMS/1294

Massimiliano Gubinelli · Herbert Koch · Tadahiro Oh

Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity

Received September 15, 2020

Abstract. Using ideas from paracontrolled calculus, we prove local well-posedness of a renormalized version of the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity forced by an additive space-time white noise on a periodic domain. There are two new ingredients as compared to the parabolic setting. (i) In constructing stochastic objects, we have to carefully exploit dispersion at a multilinear level. (ii) We introduce novel random operators and leverage their regularity to overcome the lack of smoothing of usual paradifferential commutators.

Keywords. Stochastic nonlinear wave equation, nonlinear wave equation, renormalization, white noise, paracontrolled calculus

Contents

Massimiliano Gubinelli: Hausdorff Center for Mathematics and Institut für Angewandte Mathematik, Universität Bonn, 53115 Bonn, Germany; gubinelli@iam.uni-bonn.de

Herbert Koch: Mathematisches Institut, Universität Bonn, 53115 Bonn, Germany; koch@math.uni-bonn.de

Tadahiro Oh: School of Mathematics, The University of Edinburgh, and The Maxwell Institute for the Mathematical Sciences, Edinburgh, EH9 3FD, UK; hiro.oh@ed.ac.uk

Mathematics Subject Classification (2020): Primary 35L71; Secondary 60H15

1. Introduction

1.1. Singular stochastic nonlinear wave equations

We continue the study of singular stochastic nonlinear wave equations (SNLW) driven by additive space-time white noise, initiated in [39]. There we studied the case of the SNLW equation with a polynomial nonlinearity on the two-dimensional torus $\mathbb{T}^2 = (\mathbb{R}/2\pi\mathbb{Z})^2$. By introducing a suitable renormalization of the nonlinearity, we developed a local-in-time existence and uniqueness theory. Global solutions on $\mathbb{T}^2$ have been obtained in [40] for the defocusing cubic nonlinearity. See also [95] for an analogous global well-posedness result on the Euclidean space $\mathbb{R}^2$. Here, we consider SNLW on the three-dimensional torus $\mathbb{T}^3 = (\mathbb{R}/2\pi\mathbb{Z})^3$ starting with the case of quadratic nonlinearity. Our aim is to provide a local well-posedness theory for the equation which formally reads

$$\begin{cases} \partial_t^2 u + (1-\Delta)u = -u^2 + \infty + \xi, \\ (u, \partial_t u)|_{t=0} = (u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3), \end{cases} \qquad (x,t) \in \mathbb{T}^3 \times \mathbb{R}_+, \tag{1.1}$$

where $\mathcal{H}^s(\mathbb{T}^3) = H^s(\mathbb{T}^3) \times H^{s-1}(\mathbb{T}^3)$ and $\xi(x,t)$ denotes a (Gaussian) space-time white noise on $\mathbb{T}^3 \times \mathbb{R}_+$ with space-time covariance given by

$$\mathbb{E}[\xi(x_1,t_1)\xi(x_2,t_2)] = \delta(x_1 - x_2)\delta(t_1 - t_2).$$

The expression $-u^2 + \infty$ denotes the renormalization of the product u^2. As we will see below, solutions to this equation are expected to be distributions of (spatial) regularity below $-1/2$.

We state our main result.

Theorem 1.1. *Given $1/4 < s < 1/2$, let $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$. Given $N \in \mathbb{N}$, let $\xi_N = \pi_N \xi$, where π_N is the frequency projector onto the spatial frequencies $\{|n| \le N\}$ defined in* (1.18) *below. Then there exists a sequence $\{\sigma_N(t)\}_{N\in\mathbb{N}}$ of time-dependent constants tending to ∞* (*see* (1.21) *below*) *such that, given a small $\varepsilon = \varepsilon(s) > 0$, the solution u_N to the renormalized SNLW*

$$\begin{cases} \partial_t^2 u_N + (1-\Delta)u_N = -u_N^2 + \sigma_N + \xi_N, \\ (u_N, \partial_t u_N)|_{t=0} = (u_0, u_1), \end{cases} \tag{1.2}$$

converges to a stochastic process $u \in C([0,T]; H^{-1/2-\varepsilon}(\mathbb{T}^3))$ almost surely, where $T = T(\omega)$ is an almost surely positive stopping time. The limit u is a local-in-time solution to (*the renormalized version of*) (1.1).

Furthermore, we will provide a description of the limiting distribution u in terms of *paracontrolled* distributions introduced in [38].

Let us comment on the need of the renormalized formulation (1.2). In the context of parabolic stochastic partial differential equations (SPDEs), the need and meaning of renormalization of SPDEs have been intensely studied and much progress has been achieved in recent years, starting with Da Prato and Debussche's strong solutions approach [24] to the dynamical Φ_2^4 model, continuing with Hairer's solution of the KPZ equation [43], the subsequent invention of regularity structures [45], and the discovery of alternative approaches such as paracontrolled distributions [38], Kupiainen's renormalization group approach [60, 61], and the approach of Otto, Weber, and coauthors [5, 83, 84].

On the one hand, the theory of regularity structures [30, 44, 45] has since grown into a complete framework [13, 14, 20, 46] which can deal with a large class of parabolic equations (in the so-called *subcritical* regime) such as the dynamical sine-Gordon model [21, 51], the generalized KPZ equation used to describe a natural random evolution on the space of paths over a manifold [47], and other interesting models like those related to Abelian gauge theories [88] or some equations in the full space [48]. On the other hand, the theory of paracontrolled distributions has revealed itself as an effective method for a restricted class of singular SPDEs [2, 4, 5, 17, 19, 31, 37, 41, 55–57, 67, 85, 96]. Let us also mention that certain quasilinear parabolic equations can be considered using a natural extension of these theories [32, 34, 84].

Singular SPDEs have been shown to describe large scale behavior of many random dynamical models, including both continuous [33, 50, 52–54, 90] and discrete ones [18, 49, 64, 65, 68, 89]. This phenomenon has been named *weak universality*.

Renormalization can be, in the first instance, justified by the need to obtain nontrivial (i.e. nonlinear) limiting problems. At a deeper level, singular PDEs and the need of their renormalization are tightly linked with the phenomenon of weak universality. These equations are meant to describe large scale fluctuations of well-behaved smooth random systems and, in this perspective, both the distributional nature of the solution and the renormalization have clear physical meanings: irregularity of solutions is a manifestation of microscopic random fluctuations, while renormalization is linked to the fine tuning of the parameters needed to allow for nonlinear fluctuations at the macroscopic level. While this discussion is quite informal and general, this picture can be understood rigorously in many specific cases, at least in the parabolic setting.

As far as wave equations are concerned, it has been observed in [1, 72, 73, 87] that SNLW with regularized additive space-time noise converges to a linear equation as the regularization is removed, essentially independently of the kind of (Lipschitz) nonlinearity considered. This hints that wave equations also need a certain fine tuning of the parameters in order to exhibit singular nonlinear fluctuations.

All the theories mentioned above are, however, designed to handle parabolic equations and it is not a priori clear how to adapt them to dispersive or hyperbolic phenomena.

Schrödinger and wave equations in two and three dimensions with multiplicative spatial white noise have been considered using spectral methods in [25, 26, 42]. The spatial nature of the noise allowed the authors to use techniques similar to the parabolic setting [2]. In our paper [39], we gave the first example of (nontrivial) weak universality in wave equations by showing that the renormalized SNLW on $\mathbb{T}^2$ describes a particular large scale limit of a random nonlinear wave equation with smooth noise. There, it was shown how, despite the hyperbolic setting, renormalization proceeds in a way quite similar to the parabolic case.

In the present paper, we will also show that, despite the fact that notions such as *homogeneity* (fundamental in the theory of regularity structures) or Besov–Hölder regularity (similarly fundamental in the theory of paracontrolled distributions in the parabolic setting) are less compelling in the hyperbolic setting, we can set up a paracontrolled analysis of the SNLW equation (1.1) which takes into account multilinear dispersive regularization and renormalization of resonant stochastic terms via the introduction of certain random operators, replacing the commutators that are standard in the parabolic paracontrolled approach of [38]. Let us note that the control of certain random operators already appeared in the analysis of discrete approximations to the KPZ equation in [41].

As an application of our results, we can identify the solutions to the SNLW equation (1.1) constructed in Theorem 1.1 as the universal limit of solutions to a certain class of random wave equations. Consider the following stochastic nonlinear wave equation on $(\kappa^{-1}\mathbb{T})^3 \times \mathbb{R}_+$:

$$\begin{cases} \partial_t^2 w_\kappa + (1-\Delta) w_\kappa = f(w_\kappa) + a_\kappa^{(0)} + a_\kappa^{(1)} w_\kappa + \kappa^2 \eta_\kappa, \\ (w_\kappa, \partial_t w_\kappa)|_{t=0} = (0,0), \end{cases} \tag{1.3}$$

where $\kappa > 0$, $f : \mathbb{R} \to \mathbb{R}$ is an arbitrary bounded C^3-function with bounded derivatives and η_κ is a Gaussian noise which is white in time (for simplicity) and smooth in space with finite range translation-invariant correlations (see (7.1) below for a precise definition). Here, $a_\kappa^{(0)}$ and $a_\kappa^{(1)}$ are parameters to be chosen later. It is not difficult to show that this equation has global smooth solutions. We think of this equation as a *microscopic* model of a given space-time random field w_κ living on a large spatial domain $(\kappa^{-1}\mathbb{T})^3$ and subject to a very small random driving force of order $\kappa^2 \ll 1$. For technical reasons we prefer to work in a bounded domain but the reader should think that the equation is set up in the full space and that the parameter κ sets the size of the random perturbation. In order to focus on the large scale / long time behavior of the solutions to this equation, we perform a hyperbolic rescaling of the independent variables (x,t) and introduce a new random field u_κ given by

$$u_\kappa(x,t) = \kappa^{-2} w_\kappa(\kappa^{-1}x, \kappa^{-1}t), \quad (x,t) \in \mathbb{T}^3 \times \mathbb{R}_+. \tag{1.4}$$

The following theorem gives a precise description of the limiting behavior of u_κ as $\kappa \to 0$ and as the parameters $a_\kappa^{(0)}, a_\kappa^{(1)}$ are tuned in order to have

$$f(w_\kappa) + a_\kappa^{(0)} + a_\kappa^{(1)} w_\kappa \simeq w_\kappa^2,$$

implying that $w_\kappa = 0$ is a solution of the unperturbed dynamics. Accordingly, the initial data in (1.3) are set to zero in order not to interfere with the analysis of the long time effect of the random perturbation.

Theorem 1.2. *There exists a (time-dependent) choice of the coefficients $a_\kappa^{(0)}, a_\kappa^{(1)} = O(1)$ and an almost surely positive random time T such that the random field u_κ defined in (1.4) converges in probability to a well defined limit u in $C([0,T]; H^{-1/2-\varepsilon}(\mathbb{T}^3))$ as $\kappa \to 0$. The limiting random field u is (modulo a possible rescaling) a local-in-time solution to the renormalized quadratic SNLW (1.1) with zero initial data.*

In fact, we will choose the coefficient $a_\kappa^{(1)}$ depending only on f and $\kappa > 0$, that is, deterministic and independent of time (see (7.7) below).

Remark 1.3. Equation (1.1) indeed corresponds to the stochastic nonlinear Klein–Gordon equation. The same results with inessential modifications also hold for the stochastic nonlinear wave equation, where we replace the left-hand side in (1.1) by $\partial_t^2 u - \Delta u$. In the following, we simply refer to (1.1) as the *stochastic nonlinear wave equation*.

1.2. The Da Prato–Debussche trick

Let us now describe the strategy which we used in [39, 40] to tackle the renormalization of the two-dimensional SNLW equation

$$\partial_t^2 u + (1 - \Delta)u = -u^k + \xi, \quad (x, t) \in \mathbb{T}^2 \times \mathbb{R}_+,$$

for a generic monomial nonlinearity u^k. The first step is to introduce a new variable

$$u = \Psi + v, \tag{1.5}$$

where Ψ is the stochastic convolution given by

$$\Psi(t) := \mathcal{I}\xi(t) = \int_0^t \frac{\sin((t - t')\langle\nabla\rangle)}{\langle\nabla\rangle} dW(t').$$

Here, W is a cylindrical Wiener process on $L^2(\mathbb{T}^2)$, and $\mathcal{I} = (\partial_t^2 + 1 - \Delta)^{-1}$ is the Duhamel integral operator, corresponding to the forward fundamental solution to the linear wave equation, and $\langle\nabla\rangle$ is the Fourier multiplier operator corresponding to the multiplier $\langle n\rangle = (1 + |n|^2)^{1/2}$. By a standard argument, it is easy to see that the stochastic convolution Ψ almost surely has $C(\mathbb{R}_+; W^{-\varepsilon,\infty}(\mathbb{T}^2))$ regularity for all $\varepsilon > 0$. Moreover, it can be shown that for each $t > 0$, $\Psi(t) \notin L^2(\mathbb{T}^2)$ almost surely, so that making sense of powers Ψ^k and a fortiori of the full nonlinearity u^k is an issue. The appropriate renormalization corresponds to replacing the powers $\Psi(t)^k$ by the Wick powers $:\Psi(t)^k:$ of the stochastic convolution. It then follows that the equation for the residual term $v = u - \Psi$ takes the form

$$(\partial_t^2 + 1 - \Delta)v = -\sum_{\ell=0}^{k} \binom{k}{\ell} :\Psi^\ell: v^{k-\ell}. \tag{1.6}$$

By viewing $(u_0, u_1, \Psi, :\Psi^2:, \dots, :\Psi^k:)$ as a given *enhanced* data set, we studied the fixed point problem (1.6) for v via the Strichartz estimates[1] (see Lemma 2.4 below) and we proved that the renormalized SNLW on $\mathbb{T}^2$ is locally well-posed for any integer $k \geq 2$ and is globally well-posed when $k = 3$. See also [76, 80] for a related problem for the deterministic (renormalized) NLW with random initial data.

Remark 1.4. (i) In the field of stochastic parabolic PDEs, the decomposition (1.5) is usually referred to as the Da Prato–Debussche trick [23, 24]. Note that such an idea also appears in McKean [66] and Bourgain [11] in the context of (deterministic) dispersive PDEs with random initial data, both preceding [23]. See also Burq–Tzvetkov [15].

[1] In fact, one may prove local well-posedness of (1.6) on $\mathbb{T}^2$ by Sobolev's inequality, i.e. without the Strichartz estimates; see [40].

(ii) While Ψ is not a function, it turns out that the residual part v is a function of positive regularity. Namely, the decomposition (1.5) shows that the solution u "behaves like" the stochastic convolution in the high-frequency regime (or equivalently on small scales).

For our problem on the three-dimensional torus $\mathbb{T}^3$, the Da Prato–Debussche trick does not suffice. Indeed, the stochastic convolution Ψ is less regular in three dimensions: $\Psi \in C(\mathbb{R}_+; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely for any $\varepsilon > 0$. See Lemma 3.1 below. This worse behavior also causes the higher Wick powers $:\Psi^k:$ of Ψ to become less and less regular. Correspondingly, the Cauchy problem with higher powers of the nonlinear term becomes more and more difficult to study. This is why, in this paper, we limit ourselves to the first nontrivial situation, namely the case $k = 2$ which is already not amenable to be harnessed by the Da Prato–Debussche trick. The main difficulty here is the lack of sufficient regularity for the residual term v in order for the product $\Psi \cdot v$ to be well defined. In the next subsection, we will describe in detail the difficulty and the strategy to overcome it. In particular, we will use ideas from paracontrolled calculus introduced by the first author (with Imkeller and Perkowski) [19, 38] and rewrite equation (1.1) in an appropriate form, where the residual term v is further decomposed and analyzed to expose other multilinear stochastic objects of the stochastic convolution Ψ which will be subsequently estimated via probabilistic methods (and via a detailed analysis exploiting their multilinear dispersive structures).[2]

For further reference, let us now describe the construction of Ψ in the three-dimensional setting. Let W denote a cylindrical Wiener process on $L^2(\mathbb{T}^3)$.[3] More precisely, by letting

$$e_n(x) = e^{in\cdot x}, \quad \Lambda = \bigcup_{j=0}^{2} \mathbb{Z}^j \times \mathbb{Z}_+ \times \{0\}^{2-j}, \quad \text{and} \quad \Lambda_0 = \Lambda \cup \{(0,0,0)\}, \tag{1.7}$$

we have[4]

$$\begin{aligned} W(t) &= \sum_{n\in\mathbb{Z}^3} \beta_n(t) e_n \\ &= \beta_0(t)e_0 + \sum_{n\in\Lambda} \big[\sqrt{2}\,\mathrm{Re}(\beta_n(t))\cdot\sqrt{2}\cos(n\cdot x) - \sqrt{2}\,\mathrm{Im}(\beta_n(t))\cdot\sqrt{2}\sin(n\cdot x)\big], \end{aligned} \tag{1.8}$$

---

[2]We also mention a recent preprint [27] by Deng, Nahmod, and Yue which appeared about one year after our current paper. This remarkable work elaborates the ideas introduced by Gubinelli, Imkeller and Perkowski [38] and by Bringmann [12], where the basic objects are (frequency-localized) random *nonhomogeneous* linear solutions with random potentials, thus incorporating the bad part of a nonlinearity, and introduces the so-called random averaging operators, which are well adapted to the dispersive setting. See also a very recent paper [28] by the same authors which extends the method of random averaging operators in [27] and introduces the theory of random tensors.

[3]By convention, we endow $\mathbb{T}^3$ with the normalized Lebesgue measure $(2\pi)^{-3}dx$.

[4]Note that $\{e_0, \sqrt{2}\cos(n\cdot x), \sqrt{2}\sin(n\cdot x) : n \in \Lambda\}$ forms an orthonormal basis of $L^2(\mathbb{T}^3)$ (endowed with the normalized Lebesgue measure) in the real-valued setting.

where $\{\beta_n\}_{n\in\Lambda_0}$ is a family of mutually independent complex-valued Brownian motions[5] on a fixed probability space $(\Omega, \mathcal{F}, P)$ and $\beta_{-n} := \overline{\beta_n}$ for $n \in \Lambda_0$. It is easy to see that W almost surely lies in $C^\alpha(\mathbb{R}_+; W^{-3/2-\varepsilon,\infty}(\mathbb{T}^3))$ for any $\alpha < 1/2$ and $\varepsilon > 0$. We then define the stochastic convolution Ψ in the three-dimensional setting by

$$\Psi(t) := \mathcal{I}\xi(t) = \sum_{n\in\mathbb{Z}^3} e_n \int_0^t \frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle}\, d\beta_n(t'). \tag{1.9}$$

See Lemma 3.1 below for the regularity property of Ψ.

1.3. The paracontrolled approach

In the field of stochastic parabolic PDEs, there has been significant progress over the last five years. In [45], Hairer introduced the theory of regularity structures and gave a precise meaning to certain (subcritical) singular stochastic parabolic PDEs, which are classically ill-posed. In particular, he showed that the stochastic quantization equation (SQE) on $\mathbb{T}^3$,

$$\partial_t u - \Delta u = -u^3 + \infty \cdot u + \xi, \tag{1.10}$$

is locally well-posed in an appropriate sense.

In [19], Catellier and Chouk proved an analogous local well-posedness result for SQE (1.10) via the paracontrolled calculus approach of Imkeller, Perkowski, and the first author [38]. This result was extended to global well-posedness on the torus (with a uniform-in-time bound) in a recent work [67] by Mourrat and Weber. More recently, Hofmanová and the first author [37] proved global existence of unique solutions to (1.11) on $\mathbb{R}^3$.

In [19, 45], the "solution" u to (1.10) is constructed as a unique limit of the following smoothed equation:

$$\partial_t u_\delta - \Delta u_\delta = -u_\delta^3 + C_\delta u_\delta + \xi_\delta, \tag{1.11}$$

where $\xi_\delta = \rho_\delta * \xi$ denotes the noise smoothed by a mollifier ρ_δ.[6] Here, uniqueness refers to the following: while the diverging constant C_δ depends on the choice of the mollifier ρ_δ, the limit u is independent of it. As it is written, one may wonder if u actually solves any equation in the end. In fact, one can introduce a decomposition of u analogous to (1.5) such that the residual terms satisfy a system of PDEs in the pathwise sense.

In the following, we briefly describe this decomposition of u in the paracontrolled setting. For this purpose, let us define the stochastic convolution Υ by

$$\Upsilon = (\partial_t - \Delta)^{-1}\xi.$$

[5]Here, we take β_0 to be real-valued. Moreover, we normalize β_n so that $\mathrm{Var}(\beta_n(t)) = t$. In particular, $\mathrm{Var}(\mathrm{Re}\,\beta_n(t)) = \mathrm{Var}(\mathrm{Im}\,\beta_n(t)) = t/2$ for $n \neq 0$.

[6]In [45], the mollifier ρ_δ is on both spatial and temporal variables, while it is only on spatial variables in [19]. In [60], the author employs a different kind of regularization.

Here, we have adopted Hairer's convention to denote stochastic terms by trees; the vertex "$\bullet$" in $\uparrow$ corresponds to the space-time white noise ξ, while the edge denotes the Duhamel integral operator $(\partial_t - \Delta)^{-1}$. On $\mathbb{T}^3$, $\uparrow$ has spatial regularity[7] $-\frac{1}{2}-$ and hence its powers do not make sense. Denoting the renormalized cubic power "$\uparrow^3$" by Ψ,[8] we define

$$\Upsilon = (\partial_t - \Delta)^{-1} \Psi. \tag{1.12}$$

Thanks to the parabolic smoothing of degree 2, it can be seen that Υ has regularity $\frac{1}{2}- = 3(-\frac{1}{2}-) + 2$ (see for example [69]). We now write u as

$$u = \uparrow - \Upsilon + v, \tag{1.13}$$

where v is expected to be smoother than Υ. As mentioned in Remark 1.4, the decomposition $u = \uparrow + v$ in the Da Prato–Debussche trick postulates that u behaves like $\uparrow$ on small scales. This new decomposition (1.13) postulates that the second order fluctuation of u is given by $-\Upsilon$. By further splitting v as $v = X + Y$ and introducing more stochastic objects (corresponding to Step (i) in (1.14) below), one arrives at a system of PDEs for X and Y.[9] Note that the stochastic objects thus introduced satisfy certain regularity properties in an almost sure manner. Hence, by simply viewing them as given deterministic data, we solve the resulting system for X and Y in a purely *deterministic* manner (corresponding to Step (ii) in (1.14)).

The following diagram summarizes the discussion above:

$$(u_0, \xi) \overset{\text{(i)}}{\longmapsto} \underbrace{(u_0, \uparrow, \vee, \Upsilon, \Upsilon_{\vee}, \Psi_{\vee}, \vee_{\vee})}_{\text{enhanced data set}} \overset{\text{(ii)}}{\longmapsto} (X, Y) \mapsto u = \uparrow - \Upsilon + X + Y, \tag{1.14}$$

where the ill-posed solution map $(u_0, \xi) \mapsto u$ is factorized into two steps: (i) a canonical lift, generating an enhanced data set, and (ii) a deterministic continuous solution map called the Itô–Lyons map. Note that stochastic analysis is needed only in Step (i). The decomposition (1.14) together with the equations satisfied by X and Y provides a precise meaning to the limiting equation (1.10).[10] See a nice exposition in the introduction of [67].

[7] Hereafter, we use $a-$ (and $a+$) to denote $a - \varepsilon$ (and $a + \varepsilon$, respectively) for arbitrarily small $\varepsilon > 0$. If this notation appears in an estimate, then an implicit constant is allowed to depend on $\varepsilon > 0$ (and it usually diverges as $\varepsilon \to 0$).

[8] In the three-dimensional case, it is known that the "renormalized" cubic power Ψ does not quite make sense as a distribution-valued function of time due to logarithmic divergence. Note, however, that Υ defined in (1.12) is a well-defined function.

[9] Here, we are oversimplifying the argument. In fact, this decomposition $v = X + Y$ is based on a paracontrolled ansatz, postulating that $(\partial_t - \Delta)v$ is *paracontrolled* by $\vee$. See [19, 67] for further details. We will describe the details of this step in studying SNLW; see (1.25) and (1.26).

[10] The term $\infty \cdot u$ in (1.10) is introduced so that all the terms appearing in the system for X and Y are finite. Here, ∞ is interpreted as a limit of the diverging constant C_δ in (1.11), which depends on the choice of a mollifier ρ_δ (but the limiting distribution u is independent of it). See also Remark 1.14 below.

In [45], a similar decomposition of u holds at the level of regularity structures adapted to (1.10).

In the following, we describe a procedure based on a paracontrolled ansatz. This transforms (1.1) into a system of PDEs, which we can solve by standard deterministic tools.

Remark 1.5. The theory of regularity structures introduced by Hairer [45] provides a more complete framework to study singular parabolic equations than the paracontrolled calculus introduced in [38]. However, the theory of regularity structures is more rigid and we do not know at this point how to handle stochastic wave equations in high dimensions. In particular, we do not know how to lift the Duhamel integral operator $\mathcal{I}$.

Moreover, in the parabolic setting, it is easy to predict the regularity of a product. In the theory of regularity structures, this provides an intuition of a resulting homogeneity of a product of two elements in a regularity structure.[11] In the current dispersive setting, we need to exploit a multilinear smoothing property to calculate the regularity of a product of two functions (under the Duhamel integral operator) in a much more careful manner. Hence, any implementation of regularity structures to study dispersive PDEs also needs to incorporate this extra smoothing via an explicit product structure, which seems to be highly nontrivial.

We keep our discussion at a formal level and discuss spatial regularities (= differentiability) of various objects without worrying about precise spatial Sobolev spaces that they belong to. We also use the following "rules":[12]

- A product of functions of regularities s_1 and s_2 is defined if $s_1 + s_2 > 0$. When $s_1 > 0$ and $s_1 \geq s_2$, the resulting product has regularity s_2.
- A product of stochastic objects (not depending on the unknown) is always well defined, possibly with a renormalization. The product of stochastic objects of regularities s_1 and s_2 has regularity $\min(s_1, s_2, s_1 + s_2)$.

As in the case of SQE (1.10), we use $\uparrow$ to denote the stochastic convolution Ψ for the wave equation defined in (1.9):

$$\uparrow := \Psi = \mathcal{I}(\xi) = \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}\, dW(t'). \tag{1.15}$$

In this context, the vertex "$\bullet$" in $\uparrow$ corresponds to the space-time white noise ξ, while the edge denotes the Duhamel integral operator $\mathcal{I}$. Recalling that the spatial regularity $-\frac{3}{2}-$ of the space-time white noise ξ, the smoothing under $\mathcal{I}$ shows that $\uparrow$ has (spatial) regularity $-\frac{1}{2}-$; see Lemma 3.1.

Next, we define the second order stochastic term Υ by

$$\Upsilon := \mathcal{I}(\vee) = \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} \vee(t')\, dt', \tag{1.16}$$

[11] More precisely, a product of elements in a model space T of a given regularity structure (A, T, G).

[12] In the remaining part of the paper, we will justify these rules.

where $\curlyvee$ is the renormalized version of $\uparrow^2$; see [39, Proposition 2.1] and Lemma 3.1 below. This corresponds to the second term in the Picard iteration scheme for (1.1) (with zero initial data). Note that the Wick power $\curlyvee$ has regularity $-1- = 2(-\frac{1}{2}-)$. If one proceeds with "parabolic thinking",[13] one might expect that Υ has regularity

$$0- = 2\big(-\tfrac{1}{2}-\big) + 1, \tag{1.17}$$

where we gain one derivative from the Duhamel integral operator $\mathcal{I}$, in particular from $\langle\nabla\rangle^{-1}$ in (1.16). In fact, we exhibit an extra $\frac{1}{2}$-smoothing for Υ by exploiting the explicit product structure and multilinear dispersion in (1.16).

Before proceeding further, let us introduce some notations. Given $N \in \mathbb{N}$, we define the (spatial) frequency projector π_N by

$$\pi_N u := \sum_{|n|\le N} \widehat{u}(n) e_n. \tag{1.18}$$

We then define the truncated stochastic terms $\uparrow_N$ and Υ_N by

$$\uparrow_N := \pi_N \uparrow \quad \text{and} \quad \Upsilon_N := \mathcal{I}(\curlyvee_N), \tag{1.19}$$

where $\curlyvee_N$ is the Wick power defined by

$$\curlyvee_N := (\uparrow_N)^2 - \sigma_N \tag{1.20}$$

with[14]

$$\begin{aligned} \sigma_N(t) = \mathbb{E}[(\uparrow_N(x,t))^2] &= \sum_{|n|\le N} \int_0^t \left[\frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle}\right]^2 dt' \\ &= \sum_{|n|\le N} \left\{\frac{t}{2\langle n\rangle^2} - \frac{\sin(2t\langle n\rangle)}{4\langle n\rangle^3}\right\} \sim tN. \end{aligned} \tag{1.21}$$

Note that $\curlyvee = \lim_{N\to\infty} \curlyvee_N$ in $C([0,T]; W^{-1-,\infty}(\mathbb{T}^3))$ almost surely (see Lemma 3.1 below).

Proposition 1.6. *Let $T > 0$. Then Υ_N converges to Υ in $C([0,T]; W^{1/2-\varepsilon,\infty}(\mathbb{T}^3)) \cap C^1([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely for any $\varepsilon > 0$. In particular,*

$$\Upsilon \in C([0,T]; W^{1/2-\varepsilon,\infty}(\mathbb{T}^3)) \cap C^1([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))$$

almost surely for any $\varepsilon > 0$.

This proposition shows an extra $\frac{1}{2}$-smoothing for Υ as compared to (1.17). This extra smoothing results from a multilinear interaction of waves and is a manifestation of *dispersion* (at a multilinear level), which is a key difference between dispersive and parabolic

[13] Namely, if we only count the regularity of each of $\uparrow$ in $\curlyvee$ and put them together with one degree of smoothing from the Duhamel integral operator $\mathcal{I}$ *without* taking into account the product structure and the oscillatory nature of the linear wave propagator.

[14] In our spatially homogeneous setting, the variance $\sigma_N(t)$ is independent of $x \in \mathbb{T}^3$.

equations. In proving Proposition 1.6, we combine stochastic tools with multilinear dispersive analysis, in particular, carefully estimating the (nearly) time resonant and time nonresonant contributions (see Remark 3.3). In the following, we will exploit the dispersive nature of our problem in a crucial manner.

We now write u as

$$u = \mathord{\uparrow} - \Upsilon + v. \tag{1.22}$$

Then, it follows from (1.1) and (1.22) that v satisfies

$$(\partial_t^2 + 1 - \Delta)v = -(v + \mathord{\uparrow} - \Upsilon)^2 + \curlyvee = -(v - \Upsilon)^2 - 2v\mathord{\uparrow} + 2\mathord{\uparrow}\Upsilon.$$

At the second equality, we performed the Wick renormalization: $\mathord{\uparrow}^2 \rightsquigarrow \curlyvee$. The last term $\mathord{\uparrow}\Upsilon$ has regularity $-\frac{1}{2}-$, inheriting the worse regularity of $\mathord{\uparrow}$. Hence, we expect v to have regularity at most $\frac{1}{2}- = (-\frac{1}{2}-) + 1$. In particular, the product $v\mathord{\uparrow}$ is not well defined since $(\frac{1}{2}-) + (-\frac{1}{2}-) < 0$.

In order to overcome this problem, we proceed with the paracontrolled calculus. The main ingredients for the paracontrolled approach in the parabolic setting are (i) a paracontrolled ansatz and (ii) commutator estimates. For the wave equation, however, there seems to be no smoothing for a certain relevant commutator (Remark 1.17) and we need to introduce an alternative argument.

Let us first recall the definition and basic properties of paraproducts introduced by Bony [10]; see Section 2 and [3, 38] for further details; Given $j \in \mathbb{N} \cup \{0\}$, let $\mathbf{P}_j$ be the (nonhomogeneous) Littlewood–Paley projector onto the (spatial) frequencies $\{n \in \mathbb{Z}^3 : |n| \sim 2^j\}$ such that

$$f = \sum_{j=0}^{\infty} \mathbf{P}_j f.$$

Given two functions f and g on $\mathbb{T}^3$ of regularities s_1 and s_2, we write the product fg as

$$\begin{aligned} fg &= f \olessthan g + f \circleddash g + f \ogreaterthan g \\ &:= \sum_{j<k-2} \mathbf{P}_j f\, \mathbf{P}_k g + \sum_{|j-k|\le 2} \mathbf{P}_j f\, \mathbf{P}_k g + \sum_{k<j-2} \mathbf{P}_j f\, \mathbf{P}_k g. \end{aligned} \tag{1.23}$$

The first term $f \olessthan g$ (and the third term $f \ogreaterthan g$) is called the *paraproduct* of g by f (the paraproduct of f by g, respectively) and it is always well defined as a distribution of regularity $\min(s_2, s_1 + s_2)$. On the other hand, the resonant product $f \circleddash g$ is well defined in general only if $s_1 + s_2 > 0$. In the following, we also use the notation $f \mathbin{\circledcirc\!\!\geqslant} g := f \ogreaterthan g + f \circleddash g$.

As in the study of SQE on $\mathbb{T}^3$, we now introduce our paracontrolled ansatz. Namely, we suppose that $v = u - \mathord{\uparrow} + \Upsilon$ can be decomposed as

$$v = X + Y, \tag{1.24}$$

where X and Y satisfy

$$(\partial_t^2 + 1 - \Delta)X = -2(X + Y - \Upsilon) \olessthan \mathord{\uparrow}, \tag{1.25}$$

$$(\partial_t^2 + 1 - \Delta)Y = -(X + Y - \Upsilon)^2 - 2(X + Y - \Upsilon) \mathbin{\circledcirc\!\!\geqslant} \mathord{\uparrow}. \tag{1.26}$$

Furthermore, we postulate that both X and Y have positive regularities s_1 and s_2, respectively, with $0 < s_1 < s_2$.

Remark 1.7. We say that a distribution f is *paracontrolled* (by a given distribution g) if there exists f' such that

$$f = f' \olessthan g + h$$

where h is a "smoother" remainder. See [38, Definition 3.6] for a precise definition. Note, however, that the definition in [38] is given in terms of the Besov–Hölder spaces $\mathcal{C}^s = B^s_{\infty,\infty}$ and is not necessarily useful for our dispersive problem. Formally speaking, via the decomposition (1.24) with (1.25) and the regularity assumption $0 < s_1 < s_2$, we are postulating that $(\partial_t^2 + 1 - \Delta)v$ is paracontrolled by $\upharpoonleft$.

From the first equation (1.25), we see that X has regularity $\frac{1}{2}- = (-\frac{1}{2}-) + 1$. For now, let us ignore the resonant product $-2(X + Y - \curlyvee) \circledcirc \upharpoonleft$ in (1.26) and discuss the regularity of Y. Recalling that $\curlyvee$ has regularity $\frac{1}{2}-$, we see that the paraproduct $-2(X + Y - \curlyvee) \olessthan \upharpoonleft$ (with regularity $0-$) as well as $-(X + Y - \curlyvee)^2$ in (1.26) hints that Y would have regularity $1- = (0-) + 1$. This is so of course provided that we can give a meaning to the resonant product $-2(X + Y - \curlyvee) \circledcirc \upharpoonleft$. By postulating that Y has regularity at least $\frac{1}{2} + \varepsilon$ for some $\varepsilon > 0$, we see that the resonant product $Y \circledcirc \upharpoonleft$ makes sense as a distribution of regularity $s_2 + (-\frac{1}{2}-) > 0$ without any problem. Furthermore, we can make sense of the following resonant product:

$$\curlyvee\!\!\upharpoonleft := \curlyvee \circledcirc \upharpoonleft \tag{1.27}$$

as a distribution of regularity $0- = (\frac{1}{2}-) + (-\frac{1}{2}-)$ (without renormalization).

Proposition 1.8. *Let $T > 0$. Then $\curlyvee\!\!\upharpoonleft_N := \curlyvee_N \circledcirc \upharpoonleft_N$ converges to $\curlyvee\!\!\upharpoonleft$ in the space $C([0,T]; W^{-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely for any $\varepsilon > 0$. In particular,*

$$\curlyvee\!\!\upharpoonleft \in C([0,T]; W^{-\varepsilon,\infty}(\mathbb{T}^3)) \quad \textit{almost surely for any } \varepsilon > 0.$$

If one simply writes out $\curlyvee\!\!\upharpoonleft$, then there seems to be a logarithmically divergent contribution (see (4.3)). We can, however, exploit dispersion at a multilinear level as in Proposition 1.6 and show that $\curlyvee\!\!\upharpoonleft$ is indeed a well defined distribution.

Hence, it remains to give a meaning to the resonant product $X \circledcirc \upharpoonleft$. Writing equation (1.25) in the Duhamel formulation, we have

$$X = S(t)(X_0, X_1) - 2\mathcal{I}((X + Y - \curlyvee) \olessthan \upharpoonleft), \tag{1.28}$$

where $(X, \partial_t X)|_{t=0} = (X_0, X_1) \in \mathcal{H}^{s_1}(\mathbb{T}^3)$ and $S(t)$ is the propagator for the linear wave equation defined by

$$S(t)(u_0, u_1) := \cos(t\langle\nabla\rangle)u_0 + \frac{\sin(t\langle\nabla\rangle)}{\langle\nabla\rangle}u_1.$$

We need to make sense of the resonant product between $\upharpoonleft$ and each of the terms on the right-hand side of (1.28). The next lemma establishes a regularity property of the resonant product

$$Z = Z(X_0, X_1) := (S(t)(X_0, X_1)) \circledcirc \upharpoonleft.$$

Lemma 1.9. *Given $s_1 > 0$, let $(X_0, X_1) \in \mathcal{H}^{s_1}(\mathbb{T}^3)$. Then, given $T > 0$ and $\varepsilon > 0$,*

$$Z_N := (S(t)(X_0, X_1)) \ominus \upharpoonright_N$$

converges to $Z = (S(t)(X_0, X_1)) \ominus \upharpoonright$ in $C([0,T]; H^{s_1 - 1/2 - \varepsilon}(\mathbb{T}^3))$ almost surely. In particular,

$$Z = (S(t)(X_0, X_1)) \ominus \upharpoonright \in C([0,T]; H^{s_1 - 1/2 - \varepsilon}(\mathbb{T}^3)) \quad \textit{almost surely for any } \varepsilon > 0.$$

See Section 5 for the proof.

Remark 1.10. While the proof of Lemma 1.9 is a straightforward application of the Wiener chaos estimate (Lemma 2.5), we point out that the set of probability 1 on which the conclusion of Lemma 1.9 holds depends on the choice of deterministic initial data $(X_0, X_1) \in \mathcal{H}^{s_1}(\mathbb{T}^3)$. This is analogous to the situation for the recent study of nonlinear dispersive PDEs with randomized initial data [7, 8, 15, 16, 63, 77, 78, 86], where a set of probability 1 for local-in-time or global-in-time well-posedness depends on the choice of deterministic initial data (to which a randomization is applied). See [16] for a further discussion.

The main difficulty arises in making sense of the resonant product of the second term on the right-hand side of (1.28) and $\upharpoonright$. In the parabolic setting, it is at this step that one would introduce commutators in (1.28) and exploit their smoothing properties. For our dispersive problem, however, such an argument does not seem to work (see Remark 1.17 below). This is where our discussion diverges from the parabolic case.

The main idea is to study the following *paracontrolled operator* $\mathfrak{I}_{\olessthan}$ and exhibit some smoothing property. Given a function $w \in C(\mathbb{R}_+; H^{s_1}(\mathbb{T}^3))$ with $0 < s_1 < 1/2$, define

$$\begin{aligned} \mathfrak{I}_{\olessthan}(w)(t) &:= \mathcal{I}(w \olessthan \upharpoonright)(t) = \sum_{j < k-2} \mathcal{I}(\mathbf{P}_j w \cdot \mathbf{P}_k \upharpoonright) \\ &= \sum_{n \in \mathbb{Z}^3} e_n \sum_{\substack{n_1 + n_2 = n \\ |n_1| \ll |n_2|}} \int_0^t \frac{\sin((t-t')\langle n \rangle)}{\langle n \rangle} \widehat{w}(n_1, t') \widehat{\upharpoonright}(n_2, t')\, dt'. \end{aligned} \tag{1.29}$$

Here, $|n_1| \ll |n_2|$ signifies the paraproduct $\olessthan$ in the definition of $\mathfrak{I}_{\olessthan}$.[15] In the following, we decompose the paracontrolled operator $\mathfrak{I}_{\olessthan}$ into two pieces and study them separately.

Fix small $\theta > 0$. Denoting by n_1 and n_2 the spatial frequencies of w and $\upharpoonright$ as in (1.29), we further define $\mathfrak{I}_{\olessthan}^{(1)}$ and $\mathfrak{I}_{\olessthan}^{(2)}$ as the restrictions of $\mathfrak{I}_{\olessthan}$ to $\{|n_1| \gtrsim |n_2|^\theta\}$ and $\{|n_1| \ll |n_2|^\theta\}$. More concretely, we set

$$\mathfrak{I}_{\olessthan}^{(1)}(w)(t) := \sum_{n \in \mathbb{Z}^3} e_n \sum_{\substack{n_1 + n_2 = n \\ |n_2|^\theta \lesssim |n_1| \ll |n_2|}} \int_0^t \frac{\sin((t-t')\langle n \rangle)}{\langle n \rangle} \widehat{w}(n_1, t') \widehat{\upharpoonright}(n_2, t')\, dt' \tag{1.30}$$

[15] For simplicity of presentation, we use the less precise definitions of paracontrolled operators in the remaining part of this introduction. See (5.2), (5.6), and (5.7) for the precise definitions of the paracontrolled operators $\mathfrak{I}_{\olessthan}^{(1)}$ and $\mathfrak{I}_{\olessthan, \ominus}$.

and $\mathfrak{I}^{(2)}_{\otimes}(w) := \mathfrak{I}_{\otimes}(w) - \mathfrak{I}^{(1)}_{\otimes}(w)$. As for the first paracontrolled operator $\mathfrak{I}^{(1)}_{\otimes}$, thanks to the lower bound $|n_1| \gtrsim |n_2|^\theta$ and the positive regularity of w, we exhibit some smoothing property entailing that the resonant product $\mathfrak{I}^{(1)}_{\otimes}(X + Y - \Upsilon) \circleddash \uparrow$ is well defined; see Lemma 5.1 and Corollary 5.2.

Next, we discuss the second paracontrolled operator $\mathfrak{I}^{(2)}_{\otimes}$. Our goal is to make sense of the resonant product $\mathfrak{I}^{(2)}_{\otimes}(w) \circleddash \uparrow$ for w with spatial regularity $\frac{1}{2}-$. Unlike $\mathfrak{I}^{(1)}_{\otimes}$, the operator $\mathfrak{I}^{(2)}_{\otimes}$ does not seem to possess a smoothing property and thus we need to directly study the operator $\mathfrak{I}_{\otimes,\circleddash}$ defined by

$$\begin{aligned}\mathfrak{I}_{\otimes,\circleddash}(w)(t) &:= \mathfrak{I}^{(2)}_{\otimes}(w) \circleddash \uparrow(t)\\ &= \sum_{n\in\mathbb{Z}^3} e_n \int_0^t \sum_{n_1\in\mathbb{Z}^3} \widehat{w}(n_1,t')\mathcal{A}_{n,n_1}(t,t')\,dt',\end{aligned} \tag{1.31}$$

where $\mathcal{A}_{n,n_1}(t,t')$ is given by

$$\mathcal{A}_{n,n_1}(t,t') = \mathbf{1}_{[0,t]}(t') \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|}} \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle} \widehat{\uparrow}(n_2,t')\,\widehat{\uparrow}(n_3,t). \tag{1.32}$$

Here, the condition $|n_1+n_2| \sim |n_3|$ is used to denote the spectral multiplier corresponding to the resonant product $\circleddash$ in (1.31). See (5.7) for a more precise definition.

Given $n \in \mathbb{Z}^3$ and $0 \le t_2 \le t_1$, define $\sigma_n(t_1,t_2)$ by

$$\begin{aligned}\sigma_n(t_1,t_2) &:= \mathbb{E}[\widehat{\uparrow}(n,t_1)\,\widehat{\uparrow}(-n,t_2)] = \int_0^{t_2} \frac{\sin((t_1-t')\langle n\rangle)}{\langle n\rangle}\frac{\sin((t_2-t')\langle n\rangle)}{\langle n\rangle}\,dt'\\ &= \frac{\cos((t_1-t_2)\langle n\rangle)}{2\langle n\rangle^2}t_2 + \frac{\sin((t_1-t_2)\langle n\rangle)}{4\langle n\rangle^3} - \frac{\sin((t_1+t_2)\langle n\rangle)}{4\langle n\rangle^3}.\end{aligned} \tag{1.33}$$

Recall from the definition (1.15) (see also (1.9)) that $\widehat{\uparrow}(n_2,t')$ and $\widehat{\uparrow}(n_3,t)$ are uncorrelated unless $n_2+n_3=0$, i.e. $n=n_1$. This leads to the following decomposition of $\mathcal{A}_{n,n_1}$:

$$\begin{aligned}\mathcal{A}_{n,n_1}(t,t') &= \mathbf{1}_{[0,t]}(t') \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|}} \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle}\\ &\qquad\times\big(\widehat{\uparrow}(n_2,t')\,\widehat{\uparrow}(n_3,t) - \mathbf{1}_{n_2+n_3=0}\cdot\sigma_{n_2}(t,t')\big)\\ &\quad + \mathbf{1}_{[0,t]}(t')\cdot\mathbf{1}_{n=n_1}\cdot \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\sigma_{n_2}(t,t')\\ &=: \mathcal{A}^{(1)}_{n,n_1}(t,t') + \mathcal{A}^{(2)}_{n,n_1}(t,t').\end{aligned} \tag{1.34}$$

The second term $\mathcal{A}^{(2)}_{n,n_1}$ is a (deterministic) "counter term" for the contribution in (1.32) from $n_2+n_3=0$. For this term, the condition $|n_1+n_2| \sim |n_3|$ reduces to $|n+n_2| \sim |n_2|$, which is automatically satisfied under $|n| \ll |n_2|^\theta$ for small $\theta > 0$ (see (4.11) and (4.12) below).

In view of (1.33), the sum in n_2 for the second term $\mathcal{A}^{(2)}_{n,n_1}$ is not absolutely convergent. Nonetheless, by exploiting dispersion, we show the following boundedness property of the paracontrolled operator $\mathfrak{I}_{\textcircled{<},\textcircled{=}}$ defined in (1.31). Given Banach spaces B_1 and B_2, we use $\mathcal{L}(B_1; B_2)$ to denote the space of bounded linear operators from B_1 to B_2.

Proposition 1.11. *Let $s_2 < 1$ and $T > 0$. Then there exists small $\theta = \theta(s_2) > 0$ and $\varepsilon > 0$ such that the paracontrolled operator $\mathfrak{I}_{\textcircled{<},\textcircled{=}}$ defined in (1.31) belongs to the class*

$$\mathcal{L}_1 = \mathcal{L}\big(C([0,T]; L^2(\mathbb{T}^3)) \cap C^1([0,T]; H^{-1-\varepsilon}(\mathbb{T}^3))\,;\, C([0,T]; H^{s_2-1}(\mathbb{T}^3))\big) \tag{1.35}$$

almost surely. Moreover, if we define the paracontrolled operator $\mathfrak{I}^N_{\textcircled{<},\textcircled{=}}$, $N \in \mathbb{N}$, by replacing $\uparrow$ in (1.31) and (1.32) with the truncated stochastic convolution $\uparrow_N$ of (1.19), then the truncated paracontrolled operators $\mathfrak{I}^N_{\textcircled{<},\textcircled{=}}$ converge almost surely to $\mathfrak{I}_{\textcircled{<},\textcircled{=}}$ in $\mathcal{L}_1$.

As in the proofs of Propositions 1.6 and 1.8, dispersion plays an essential role in establishing the regularity property of the paracontrolled operator $\mathfrak{I}_{\textcircled{<},\textcircled{=}}$. See Section 5 for the proof.

Putting all together, we obtain the following system of PDEs for X and Y:

$$\begin{aligned}
(\partial_t^2 + 1 - \Delta)X &= -2(X + Y - \curlyvee) \mathbin{\textcircled{<}} \uparrow,\\
(\partial_t^2 + 1 - \Delta)Y &= -(X + Y - \curlyvee)^2 - 2(X + Y - \curlyvee) \mathbin{\textcircled{>}} \uparrow\\
&\quad - 2Y \mathbin{\textcircled{=}} \uparrow + 2\curlyvee^{\bullet}_{\textcircled{=}} - 2Z\\
&\quad + 4\mathfrak{I}^{(1)}_{\textcircled{<}}(X + Y - \curlyvee) \mathbin{\textcircled{=}} \uparrow + 4\mathfrak{I}_{\textcircled{<},\textcircled{=}}(X + Y - \curlyvee),\\
(X, \partial_t X, Y, \partial_t Y)|_{t=0} &= (X_0, X_1, Y_0, Y_1).
\end{aligned} \tag{1.36}$$

Let $s_1 < 1/2$ and fix a pair of deterministic functions (X_0, X_1) in $\mathcal{H}^{s_1}(\mathbb{T}^3)$. The stochastic terms and operator appearing in the system (1.36) are

$$\uparrow, \quad \curlyvee, \quad \curlyvee^{\bullet}_{\textcircled{=}}, \quad Z = Z(X_0, X_1), \quad \text{and} \quad \mathfrak{I}_{\textcircled{<},\textcircled{=}}, \tag{1.37}$$

In Lemma 3.1 and Propositions 1.6 and 1.8, we study the regularity properties of $\uparrow$, $\curlyvee$, and $\curlyvee^{\bullet}_{\textcircled{=}}$, and show that each of these terms belongs almost surely to $C(\mathbb{R}_+; W^{s,\infty}(\mathbb{T}^3))$ with the regularity s shown in Table 1. In Lemma 1.9, we prove that $Z \in C(\mathbb{R}_+; H^s(\mathbb{T}^3))$ almost surely for $s < s_1 - 1/2$. In Proposition 1.11, we establish the almost sure boundedness property of the paracontrolled operator $\mathfrak{I}_{\textcircled{<},\textcircled{=}}$ in an appropriate space. We summarize these regularity properties in Table 1.

	$\uparrow$	$\curlyvee$	$\curlyvee^{\bullet}_{\textcircled{=}}$	Z	$\mathfrak{I}_{\textcircled{<},\textcircled{=}}$
s	$-1/2 - \varepsilon$	$1/2 - \varepsilon$	$-\varepsilon$	$s_1 - 1/2 - \varepsilon$	$\mathcal{L}_1$ in (1.35)

Tab. 1. The list of relevant stochastic terms with their regularities.

In Lemma 5.1 and Corollary 5.2, we also study the regularity property of the paracontrolled operator $\mathfrak{I}^{(1)}_{\textcircled{<}}$.

We now state our main result on local well-posedness of the system (1.36), viewing the terms and operators in (1.37) as *predefined deterministic* data with certain regularity properties.

Theorem 1.12. *Let* $1/4 < s_1 < 1/2 < s_2 \leq s_1 + 1/4$*. There exist small* $\theta = \theta(s_2) > 0$ *and* $\varepsilon = \varepsilon(s_1, s_2, \theta) > 0$ *such that if*

- $\upharpoonleft$*,* $\curlyvee$*, and* $\curlyvee^{\!\!\bullet}_{\circledequal}$ *are distributions belonging to* $C(\mathbb{R}_+; W^{s,\infty}(\mathbb{T}^3))$ *for* s *as in Table 1, and moreover*
$$\curlyvee \in C^1(\mathbb{R}_+; W^{-1-\varepsilon,\infty}(\mathbb{T}^3)),$$
- Z *is a distribution belonging to* $C(\mathbb{R}_+; H^{s_1-1/2-\varepsilon}(\mathbb{T}^3))$*,*
- *the operator* $\mathfrak{I}_{\olessthan,\circledequal}$ *belongs to the class* $\mathcal{L}_1$ *of* (1.35)*,*

then the system (1.36) *is locally well-posed in* $\mathcal{H}^{s_1}(\mathbb{T}^3) \times \mathcal{H}^{s_2}(\mathbb{T}^3)$*. More precisely, given any* $(X_0, X_1, Y_0, Y_1) \in \mathcal{H}^{s_1}(\mathbb{T}^3) \times \mathcal{H}^{s_2}(\mathbb{T}^3)$*, there exists* $T > 0$ *such that there exists a unique solution* (X, Y) *to the system* (1.36) *on* $[0, T]$ *in the class*

$$\begin{aligned} Z_T^{s_1,s_2} &= X_T^{s_1} \times Y_T^{s_2} \\ &\subset C([0,T]; H^{s_1}(\mathbb{T}^3) \times H^{s_2}(\mathbb{T}^3)) \cap C^1([0,T]; H^{s_1-1}(\mathbb{T}^3) \times H^{s_2-1}(\mathbb{T}^3)), \end{aligned}$$

depending continuously on the enhanced data set

$$\Xi = (X_0, X_1, Y_0, Y_1, \upharpoonleft, \curlyvee, \curlyvee^{\!\!\bullet}_{\circledequal}, Z, \mathfrak{I}_{\olessthan,\circledequal}) \tag{1.38}$$

in the class

$$\begin{aligned} \mathcal{X}_T^{s_1,s_2,\varepsilon} &= \mathcal{H}^{s_1}(\mathbb{T}^3) \times \mathcal{H}^{s_2}(\mathbb{T}^3) \times C([0,T]; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3)) \\ &\quad \times (C([0,T]; W^{1/2-\varepsilon,\infty}(\mathbb{T}^3)) \cap C^1([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))) \\ &\quad \times C([0,T]; W^{-\varepsilon,\infty}(\mathbb{T}^3)) \times C([0,T]; H^{s_1-1/2-\varepsilon}(\mathbb{T}^3)) \times \mathcal{L}_1. \end{aligned} \tag{1.39}$$

Here, $X_T^{s_1}$ *and* $Y_T^{s_2}$ *are the energy spaces at the regularities* s_1 *and* s_2 *intersected with appropriate Strichartz spaces* (*see* (6.1) *below*).

Theorem 1.12 states local well-posedness of the system (1.36) when we view the enhanced data set $\upharpoonleft$, $\curlyvee$, $\curlyvee^{\!\!\bullet}_{\circledequal}$, Z, and $\mathfrak{I}_{\olessthan,\circledequal}$ as given deterministic distributions or operator. Hence, the proof of Theorem 1.12 is entirely deterministic. By writing (1.36) in the Duhamel formulation

$$\begin{aligned} X(t) &= \Phi_1(X,Y)(t) \\ &:= S(t)(X_0, X_1) - 2\int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}[(X+Y-\curlyvee)\olessthan\upharpoonleft](t')\,dt', \\ Y(t) &= \Phi_2(X,Y)(t) \\ &:= S(t)(Y_0, Y_1) - \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}\big[(X+Y-\curlyvee)^2 + 2(X+Y-\curlyvee)\ogreaterthan\upharpoonleft \\ &\qquad + 2Y\circledequal\upharpoonleft - 2\curlyvee^{\!\!\bullet}_{\circledequal} + 2Z \\ &\qquad - 4\mathfrak{I}^{(1)}_{\olessthan}(X+Y-\curlyvee)\circledequal\upharpoonleft - 4\mathfrak{I}_{\olessthan,\circledequal}(X+Y-\curlyvee)\big](t')\,dt', \end{aligned} \tag{1.40}$$

we show that the map $\Phi = (\Phi_1, \Phi_2)$ is a contraction on a closed ball in $Z_T^{s_1,s_2}$ for sufficiently small $T > 0$ which depends only on the $\mathcal{X}_T^{s_1,s_2,\varepsilon}$-norm of the enhanced data set Ξ in (1.38). The main tools are (i) the Strichartz estimates for the wave equations (Lemma 2.4) and (ii) the paraproduct estimates (Lemma 2.1). See Section 6 for details.

Finally, let us discuss the consequence of Theorem 1.12 for the original SNLW (1.1).

Proof of Theorem 1.1. Let $1/4 < s < 1/2$. Given $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$, let $(X_0, X_1, Y_0, Y_1) = (u_0, u_1, 0, 0)$. For each $N \in \mathbb{N}$, we construct the enhanced data set associated with the truncated noise $\xi_N = \pi_N \xi$,

$$\Xi_N = (u_0, u_1, 0, 0, \uparrow_N, \curlyvee_N, \curlyvee^{\circ}_N, Z_N, \mathfrak{Z}^N_{\varolessthan,\circeq}).$$

Here, $\uparrow_N$, $\curlyvee_N$, $\curlyvee^{\circ}_N$, and $\mathfrak{Z}^N_{\varolessthan,\circeq}$ are as in (1.19) and in Propositions 1.8 and 1.11, while we set $Z_N = Z_N(u_0, u_1) = S(t)(u_0, u_1) \circeq \uparrow_N$. Let (X_N, Y_N) be the unique local-in-time solution to the system (1.36) with the enhanced data set Ξ_N and define u_N by

$$u_N = \uparrow_N - \curlyvee_N + X_N + Y_N. \tag{1.41}$$

Then, by reversing the discussion above with the use of (1.20), we see that u_N satisfies the renormalized SNLW (1.2) *provided* σ_N is chosen as in (1.21).

From Lemma 3.1, Propositions 1.6, 1.8, Lemma 1.9, Corollary 5.2, and Proposition 1.11, we see that Ξ_N converges almost surely to

$$\Xi = (u_0, u_1, 0, 0, \uparrow, \curlyvee, \curlyvee^{\circ}, S(t)(u_0, u_1) \circeq \uparrow, \mathfrak{Z}_{\varolessthan,\circeq}) \tag{1.42}$$

in the $\mathcal{X}_1^{s,1/2+\varepsilon,\varepsilon}$-topology for some small $\varepsilon > 0$. Then, the (pathwise) continuous dependence of the solution map for the system (1.36) on the enhanced data set in $\mathcal{X}_1^{s,1/2+\varepsilon,\varepsilon}$ implies that

- the (random) local existence time $T = T(\omega)$ depicted in Theorem 1.12 can be chosen uniformly for $\{(X_N, Y_N)\}_{N\in\mathbb{N}}$ and (X, Y). Here, (X, Y) is the unique solution to (1.36) with the enhanced data Ξ in (1.42).
- the solution u_N to the renormalized SNLW (1.2) defined in (1.41) converges almost surely to u in $C([0, T]; H^{-1/2-\varepsilon}(\mathbb{T}^3))$, where u is given by

$$u = \uparrow - \curlyvee + X + Y. \tag{1.43}$$

This proves Theorem 1.1 under the condition that σ_N is chosen as in (1.21). ∎

Remark 1.13. As we pointed out in Remark 1.10, the set Σ of probability 1 on which Theorem 1.1 holds depends on the choice of (deterministic) initial data $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$ due to Lemma 1.9. If we assume a slightly higher regularity, namely, if we work with $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$ for some $s > 1/2$, we can choose the set Σ of probability 1, independent of $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$, by simply setting $(X_0, X_1, Y_0, Y_1) = (0, 0, u_0, u_1)$, which avoids the use of Lemma 1.9.

Remark 1.14. Given $\rho \in C_c^\infty(\mathbb{R}^3)$ with $\int_{\mathbb{R}^3} \rho(x)\,dx = 1$ and $\operatorname{supp} \rho \subset [-1/2, 1/2)^3 \simeq \mathbb{T}^3$, we define a smooth mollifier ρ_δ, $0 < \delta \le 1$, by setting

$$\rho_\delta(x) = \delta^{-3}\rho(\delta^{-1}x). \tag{1.44}$$

We also say that such a ρ is a *mollification kernel*. Then, the same argument leading to Theorem 1.1 can be used to prove the following convergence and uniqueness statement (see [19,45]). Given $1/4 < s < 1/2$, let $(u_0, u_1) \in \mathcal{H}^s(\mathbb{T}^3)$. Let $\xi_\delta = \rho_\delta * \xi$ be the noise smoothed by a smooth mollifier ρ_δ. Then, for any $0 < \delta \le 1$, there exists $C_\delta = C_\delta(t, \rho)$ such that the solution u_δ to the smoothed SNLW

$$\begin{cases} \partial_t^2 u_\delta + (1 - \Delta) u_\delta = -u_\delta^2 + C_\delta + \xi_\delta, \\ (u_\delta, \partial_t u_\delta)|_{t=0} = (u_0, u_1), \end{cases}$$

converges in probability to some distribution u in $C([0, T]; H^{-1/2-\varepsilon}(\mathbb{T}^3))$ for any $\varepsilon > 0$, where $T = T(\omega)$ is an almost surely positive stopping time, independent of $0 < \delta \le 1$. Here, we have $C_\delta(t, \rho) = C_0 t/\delta + C(t, \rho)$, where C_0 is a universal constant and $C(t, \rho)$ is a finite constant. Moreover, the limit u is unique in the sense that it is independent of the choice of the mollification kernel ρ.

In the proof of Proposition 1.11 presented in Section 5 below, we make use of a certain symmetry, which may seem to suggest that the Fourier transform of the mollification kernel ρ needs to be symmetric. It is, however, possible to extend Theorem 1.1 to a general mollification even if the Fourier transform of ρ is not symmetric. See Remark 5.4 for a further discussion.

Furthermore, we point out that we can also consider space-time mollifiers and obtain an analogous result. In this case, we impose an additional assumption[16] that a space-time mollification kernel $\rho(x, t)$ is even in x, namely, $\rho(-x, t) = \rho(x, t)$ for any $t \in \mathbb{R}$ (see Remark 5.5). In the context of Theorem 1.2 on weak universality, this modification allows us to handle noises that are smooth in both space and time.

Let us complete this section by some additional observations.

Remark 1.15. As we saw in (1.21), the variance $\sigma_N(t)$ of the truncated stochastic convolution $\uparrow_N$ is time dependent, resulting in a time-dependent renormalization constant in Theorem 1.1. This is due to the lack of any dissipation mechanism in the dispersive setting. In the parabolic setting, for example in the case of SQE (1.10), there exists a unique invariant measure for the truncated linear stochastic dynamics, which allows us to take time-independent renormalization constants. In the wave equation case, we

[16] As pointed out in [43,45], in the case of the KPZ equation, regularization via a nonsymmetric space-time mollifier can cause the appearance of an additional transport term: see [30, Proposition 15.12 and Remark 15.13]. For the quadratic SNLW (1.1) on $\mathbb{T}^3$, it may also be possible to use regularization via a nonsymmetric space-time mollifier by introducing new types of counter terms for $\curlyvee$ and the paracontrolled operator $\mathfrak{I}_{\circledless,\circleddash}$. We, however, do not pursue this issue in this paper.

may consider the equation with damping, namely, replace the left-hand side of (1.1) by $\partial_t^2 u + \partial_t u + (1 - \Delta)u$ so that there exists a unique invariant Gaussian measure μ_N for the (truncated) linear dynamics. In this case, by taking the initial data distributed according to this invariant Gaussian measure μ_N, the variance of the truncated stochastic convolution becomes time independent and thus we can use a time-independent renormalization constant. See [40, 75, 79, 82].

We point out that in the parabolic setting, it is possible to start with arbitrary deterministic initial data u_0 (under some regularity assumption) and still use time-independent renormalization constants. This is thanks to strong parabolic smoothing, allowing us to handle rough initial data of the form $u_0 - z_0$, where z_0 is a random function distributed by the massive Gaussian free field:

$$z_0 = \sum_{n \in \mathbb{Z}^3} \frac{g_n}{\langle n \rangle} e_n.$$

Here, $\{g_n\}_{n \in \Lambda_0}$ is a sequence of independent standard complex-valued Gaussian random variables and $g_{-n} := \overline{g_n}$, $n \in \Lambda_0$. On the other hand, in the damped wave case, due to the lack of strong smoothing, our solution theory does not allow us to handle random data of the form $(u_0, u_1) - (z_0, z_1)$, where z_0 is as above and z_1 is distributed by the spatial white noise. Unfortunately, such initial data is too rough to handle in the deterministic manner for the damped wave equation. This in particular implies that for the damped wave equation, it is not possible to start with arbitrary deterministic initial data (under some regularity assumption) and use a time-independent renormalization constant (see also [75, Remark 1.2 (iii)].

Remark 1.16. (i) In making sense of the resonant product $X \circleddash \uparrow$, we substituted the Duhamel formula for X as in (1.28). This is analogous to the treatment of SQE (1.10); see [67]. Note that such an iteration of the (part of) Duhamel formula already appears in the study of the stochastic KdV equation with an additive (almost) white noise (see [36, 74]).

(ii) Unlike the parabolic setting, we need to assume higher regularity for the initial data than that for the stochastic convolution. This is due to the lack of smoothing in our dispersive problem. If the initial data is random (independent of the additive space-time white noise), we may take it to be of low regularity.

(iii) In Proposition 1.11, we assumed C^1-regularity in time of the input function for the paracontrolled operator $\mathfrak{I}_{\circledcirc, \circleddash}$. This smoothness in time allows us to exploit the time oscillation by integration by parts (see (5.19) below). On the one hand, we may prove an analogous boundedness result by assuming less time regularity of the input function. On the other hand, it seems that we do need to assume some time regularity of the input function. This necessity for smoothness in time is analogous to the parabolic setting, but for a different reason; see [19, 45, 67].

Remark 1.17 (On commutators). As mentioned above, commutators play an important role in applying the paracontrolled calculus in the parabolic setting. If we were to follow the argument for SQE presented in [67], we would write (1.28) as

$$\begin{aligned} X &= S(t)(X_0, X_1) - 2\mathcal{I}((X + Y + \curlyvee) \mathbin{\circledless} \uparrow) \\ &= S(t)(X_0, X_1) - 2(X + Y + \curlyvee) \mathbin{\circledless} \mathcal{I}(\uparrow) + \mathsf{com}_1. \end{aligned} \tag{1.45}$$

Here, com_1 denotes the commutator of the paraproduct $\mathbin{\circledless}$ and the Duhamel integral operator $\mathcal{I} = (\partial_t^2 + 1 - \Delta)^{-1}$.

In the case of SQE (1.10) on $\mathbb{T}^3$, it was crucial that the commutator of the paraproduct $\mathbin{\circledless}$ and the Duhamel integral operator $(\partial_t - \Delta)^{-1}$ for the heat equation enjoyed some smoothing property, which resulted from the smoothing property of the commutator $[e^{t\Delta}, \mathbin{\circledless}]$ (see [19, Lemma 2.5] and [67, Proposition A.16]). Unfortunately, in our dispersive setting, the commutator com_1 does not seem to provide any smoothing. We point out that if the identity (1.45) were to hold with a smoother commutator, then the rest would follow as in the parabolic setting [67] (and in particular, there would be no need to introduce paracontrolled operators). Namely, by defining

$$[\mathbin{\circledless}, \mathbin{\circledequal}](f, g, h) = (f \mathbin{\circledless} g) \mathbin{\circledequal} h - f(g \mathbin{\circledequal} h),$$

we can write

$$X \mathbin{\circledequal} \uparrow = S(t)(X_0, X_1) - 2(X + Y + \curlyvee)(\mathcal{I}(\uparrow) \mathbin{\circledequal} \uparrow) + \mathsf{com}_1 \mathbin{\circledequal} \uparrow + \mathsf{com}_2,$$

where $\mathsf{com}_2 = [\mathbin{\circledless}, \mathbin{\circledequal}](X + Y + \curlyvee, \mathcal{I}(\uparrow), \uparrow)$. Note that com_2 is a well defined distribution thanks to the smoothing property of $[\mathbin{\circledless}, \mathbin{\circledequal}]$. See [38, Lemma 2.4] and [67, Proposition A.9].

Let us now consider the first commutator com_1. Given an operator T, let

$$[T, \mathbin{\circledless}](f, g) = T(f \mathbin{\circledless} g) - f \mathbin{\circledless} (Tg).$$

Then, by setting $\mathcal{S} = \langle \nabla \rangle \mathcal{I} = \langle \nabla \rangle (\partial_t^2 + 1 - \Delta)^{-1}$, we have

$$[\mathcal{I}, \mathbin{\circledless}](f, g) = \mathcal{S} \circ [\langle \nabla \rangle^{-1}, \mathbin{\circledless}](f, g) + [\mathcal{S}, \mathbin{\circledless}](f, \langle \nabla \rangle^{-1} g).$$

It is easy to see that the first commutator $[\langle \nabla \rangle^{-1}, \mathbin{\circledless}]$ enjoys certain smoothing.[17] On the other hand, if we were to exhibit smoothing for the second commutator $[\mathcal{S}, \mathbin{\circledless}]$ as in the parabolic setting, we would need to study the smoothing property of the commutator

[17] If f and g have regularities $0 < s_1 < 1$ and $s_2 < 0$ with $s_1 + s_2 < 0$, then each of $\langle \nabla \rangle^{-1}(f \mathbin{\circledless} g)$ and $f \mathbin{\circledless} (\langle \nabla \rangle^{-1} g)$ has regularity $s_2 + 1$. On the other hand, the commutator $[\langle \nabla \rangle^{-1}, \mathbin{\circledless}](f, g)$ has regularity $s_1 + s_2 + 1$, which, roughly speaking, follows from the following observation; given $n, n_1, n_2 \in \mathbb{Z}^3$ with $n = n_1 + n_2$, we have

$$\left| \frac{1}{\langle n \rangle} - \frac{1}{\langle n_2 \rangle} \right| = \frac{\left| \langle n_2 \rangle - \langle n \rangle \right|}{\langle n \rangle \langle n_2 \rangle} \lesssim \frac{\langle n_1 \rangle}{\langle n \rangle \langle n_2 \rangle}.$$

In particular, when $|n_1| \ll |n_2| \sim |n|$ and the first function f has positive regularity, this observation provides smoothing.

$[\sin(t\langle\nabla\rangle), \circledcirc]$. Unfortunately, there does not seem to be any smoothing for this commutator in general,[18] which prevents us from working with commutators for our dispersive problem. By introducing paracontrolled operators, we indeed exhibit smoothing under the commutator $[\mathcal{S}, \circledcirc]$ (and hence under $[\mathcal{I}, \circledcirc]$) in a probabilistic manner with a specific second input function, i.e. $g = \uparrow$ (see Proposition 1.11). This is in sharp contrast with the parabolic setting, where a smoothing can be shown for $[e^{t\Delta}, \circledcirc]$ in a deterministic manner (without specifying the second input function).

Lastly, we point out that our approach via paracontrolled operators also works in the parabolic setting. In particular, in place of using commutators, we can study the relevant paracontrolled operators directly to prove local well-posedness of SQE (1.10) on $\mathbb{T}^3$.

2. Notations and basic lemmas

2.1. Sobolev spaces and Besov spaces

Let $s \in \mathbb{R}$ and $1 \le p \le \infty$. We define the L^2-based Sobolev space $H^s(\mathbb{T}^3)$ by the norm

$$\|f\|_{H^s} = \|\langle n\rangle^s \widehat{f}(n)\|_{\ell^2_n}$$

and set

$$\mathcal{H}^s(\mathbb{T}^3) = H^s(\mathbb{T}^3) \times H^{s-1}(\mathbb{T}^3).$$

We also define the L^p-based Sobolev space $W^{s,p}(\mathbb{T}^3)$ by the norm

$$\|f\|_{W^{s,p}} = \|\mathcal{F}^{-1}(\langle n\rangle^s \widehat{f}(n))\|_{L^p}$$

with the standard modification when $p = \infty$. When $p = 2$, we have $H^s(\mathbb{T}^3) = W^{s,2}(\mathbb{T}^3)$.

Let $\phi : \mathbb{R} \to [0, 1]$ be a smooth bump function supported on $[-8/5, 8/5]$ and $\phi \equiv 1$ on $[-5/4, 5/4]$. For $\xi \in \mathbb{R}^3$, we set $\phi_0(\xi) = \phi(|\xi|)$ and

$$\phi_j(\xi) = \phi(|\xi|/2^j) - \phi(|\xi|/2^{j-1})$$

for $j \in \mathbb{N}$. Then, for $j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, we define the Littlewood–Paley projector $\mathbf{P}_j$ as the Fourier multiplier operator with symbol

$$\varphi_j(\xi) = \frac{\phi_j(\xi)}{\sum_{k\in\mathbb{N}_0} \phi_k(\xi)}. \tag{2.1}$$

Note that, for each $\xi \in \mathbb{R}^3$, the sum in the denominator is over finitely many k's. Thanks to the normalization (2.1), we have

$$f = \sum_{j=0}^{\infty} \mathbf{P}_j f,$$

which is used in (1.23).

[18] Under $|n_1| \ll |n_2|$, there is no smoothing for $\sin(t\langle n_1 + n_2\rangle) - \sin(t\langle n_2\rangle)$.

We briefly recall the basic properties of the Besov spaces $B^s_{p,q}(\mathbb{T}^3)$ defined by the norm

$$\|u\|_{B^s_{p,q}} = \big\| 2^{sj} \|\mathbf{P}_j u\|_{L^p_x} \big\|_{\ell^q_j(\mathbb{N}_0)}.$$

Note that $H^s(\mathbb{T}^3) = B^s_{2,2}(\mathbb{T}^3)$.

Lemma 2.1. (i) (Paraproduct and resonant product estimates) *Let* $s_1, s_2 \in \mathbb{R}$ *and* $1 \le p, p_1, p_2, q \le \infty$ *be such that* $1/p = 1/p_1 + 1/p_2$. *Then*

$$\| f \otimes g\|_{B^{s_2}_{p,q}} \lesssim \|f\|_{L^{p_1}} \|g\|_{B^{s_2}_{p_2,q}}. \tag{2.2}$$

When $s_1 < 0$, *we have*

$$\| f \otimes g\|_{B^{s_1+s_2}_{p,q}} \lesssim \|f\|_{B^{s_1}_{p_1,q}} \|g\|_{B^{s_2}_{p_2,q}}. \tag{2.3}$$

When $s_1 + s_2 > 0$, *we have*

$$\| f \circledcirc g\|_{B^{s_1+s_2}_{p,q}} \lesssim \|f\|_{B^{s_1}_{p_1,q}} \|g\|_{B^{s_2}_{p_2,q}}. \tag{2.4}$$

(ii) *Let* $s_1 < s_2$ *and* $1 \le p, q \le \infty$. *Then*

$$\|u\|_{B^{s_1}_{p,q}} \lesssim \|u\|_{W^{s_2,p}}. \tag{2.5}$$

The product estimates (2.2)–(2.4) follow easily from the definition (1.23) of the paraproduct and the resonant product. See [3, 68] for details of the proofs in the nonperiodic case (which can be easily extended to the current periodic setting). The embedding (2.5) follows from the ℓ^q-summability of $\{2^{(s_1-s_2)j}\}_{j\in\mathbb{N}_0}$ for $s_1 < s_2$ and the uniform boundedness of the Littlewood–Paley projector $\mathbf{P}_j$.

We also recall the following fractional Leibniz rule.

Lemma 2.2. *Let* $0 \le s \le 1$. *Suppose that* $1 < p_j, q_j, r < \infty$, $\frac{1}{p_j} + \frac{1}{q_j} = \frac{1}{r}$, $j = 1, 2$. *Then*

$$\|\langle\nabla\rangle^s (fg)\|_{L^r(\mathbb{T}^d)} \lesssim \|f\|_{L^{p_1}(\mathbb{T}^d)} \|\langle\nabla\rangle^s g\|_{L^{q_1}(\mathbb{T}^d)} + \|\langle\nabla\rangle^s f\|_{L^{p_2}(\mathbb{T}^d)} \|g\|_{L^{q_2}(\mathbb{T}^d)}.$$

This lemma follows from the Coifman–Meyer theorem on $\mathbb{R}^d$ (see [22] and [70, inequality (1.1)]) and the transference principle [29, Theorem 3].

2.2. *On discrete convolutions*

Next, we recall the following basic lemma on a discrete convolution.

Lemma 2.3. (i) *Let* $d \ge 1$ *and* $\alpha, \beta \in \mathbb{R}$ *satisfy*

$$\alpha + \beta > d \quad \text{and} \quad \alpha, \beta < d.$$

Then

$$\sum_{n_1+n_2=n} \frac{1}{\langle n_1\rangle^\alpha \langle n_2\rangle^\beta} \lesssim \langle n\rangle^{d-\alpha-\beta} \quad \text{for any } n \in \mathbb{Z}^d.$$

(ii) *Let* $d \geq 1$ *and* $\alpha, \beta \in \mathbb{R}$ *satisfy* $\alpha + \beta > d$. *Then*

$$\sum_{\substack{n_1+n_2=n \\ |n_1|\sim|n_2|}} \frac{1}{\langle n_1\rangle^\alpha \langle n_2\rangle^\beta} \lesssim \langle n\rangle^{d-\alpha-\beta} \quad \text{for any } n \in \mathbb{Z}^d.$$

Note that, in the resonant case (ii), we do not have the restriction $\alpha, \beta < d$. Lemma 2.3 follows from elementary computations; see, for example, [69, Lemmas 4.1 and 4.2].

2.3. Strichartz estimates

Given $0 \leq s \leq 1$, we say that a pair (q, r) is *s-admissible* (a pair $(\widetilde{q}, \widetilde{r})$ is *dual s-admissible*,[19] respectively) if $1 \leq \widetilde{q} < 2 < q \leq \infty$, $1 < \widetilde{r} \leq 2 \leq r < \infty$,

$$\frac{1}{q} + \frac{3}{r} = \frac{3}{2} - s = \frac{1}{\widetilde{q}} + \frac{3}{\widetilde{r}} - 2, \quad \frac{1}{q} + \frac{1}{r} \leq \frac{1}{2}, \quad \text{and} \quad \frac{1}{\widetilde{q}} + \frac{1}{\widetilde{r}} \geq \frac{3}{2}.$$

We refer to the first two equalities as the *scaling conditions* and the last two inequalities as the *admissibility conditions*.

We say that u is a solution to the nonhomogeneous linear wave equation

$$\begin{cases} (\partial_t^2 + 1 - \Delta)u = f, \\ (u, \partial_t u)|_{t=0} = (u_0, u_1), \end{cases} \tag{2.6}$$

on a time interval containing $t = 0$ if u satisfies the following Duhamel formulation:

$$u = \cos(t\langle\nabla\rangle)u_0 + \frac{\sin(t\langle\nabla\rangle)}{\langle\nabla\rangle}u_1 + \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} f(t')\,dt'.$$

In the following, we often use the shorthand notation

$$\mathcal{I}(f)(t) = \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} f(t')\,dt'.$$

We now recall the Strichartz estimates for solutions to (2.6).

Lemma 2.4. *Given* $0 \leq s \leq 1$, *let* (q, r) *and* $(\widetilde{q}, \widetilde{r})$ *be s-admissible and dual s-admissible pairs, respectively. Then every solution u to the nonhomogeneous linear wave equation* (2.6) *satisfies, for all* $0 < T \leq 1$,

$$\|(u, \partial_t u)\|_{L_T^\infty \mathcal{H}_x^s} + \|u\|_{L_T^q L_x^r} \lesssim \|(u_0, u_1)\|_{\mathcal{H}^s} + \|f\|_{L_T^{\widetilde{q}} L_x^{\widetilde{r}}},$$
$$\|(u, \partial_t u)\|_{L_T^\infty \mathcal{H}_x^s} + \|u\|_{L_T^q L_x^r} \lesssim \|(u_0, u_1)\|_{\mathcal{H}^s} + \|f\|_{L_T^1 H_x^{s-1}}.$$

Here, we use the shorthand notation $L_T^q L_x^r = L^q([0, T]; L^r(\mathbb{T}^3))$ *etc.*

[19] Here, we define the notion of dual s-admissibility for the convenience of presentation. Note that $(\widetilde{q}, \widetilde{r})$ is dual s-admissible if and only if $(\widetilde{q}', \widetilde{r}')$ is $(1 - s)$-admissible.

The Strichartz estimates on $\mathbb{R}^d$ have been studied extensively by many mathematicians: see [35, 58, 62] in the context of the wave equation, and [59] for the Klein–Gordon equation under consideration. Thanks to the finite speed of propagation, these estimates on $\mathbb{T}^3$ follow from the corresponding estimates on $\mathbb{R}^3$.

In proving Theorem 1.12, we use the fact that $(8, \frac{8}{3})$ and $(4, 4)$ are $\frac{1}{4}$-admissible and $\frac{1}{2}$-admissible, respectively. We also use a dual $\frac{1}{2}$-admissible pair $(\frac{4}{3}, \frac{4}{3})$. In proving Theorem 1.2, we use $(\frac{4}{1+2\sigma}, \frac{4}{1-2\sigma})$ and $(\frac{4}{3+8\sigma}, \frac{4}{3-4\sigma})$ which are $(\frac{1}{2}+\sigma)$-admissible and dual $(\frac{1}{2}+\sigma)$-admissible, respectively, for small $\sigma > 0$.

2.4. *Tools from stochastic analysis*

We conclude this section by recalling useful lemmas from stochastic analysis (see [9, 91] for basic definitions). Let (H, B, μ) be an abstract Wiener space, i.e. μ is a Gaussian measure on a separable Banach space B with $H \subset B$ as its Cameron–Martin space. Given a complete orthonormal system $\{e_j\}_{j\in\mathbb{N}} \subset B^*$ of $H^* = H$, we define a *polynomial chaos* of order k to be an element of the form

$$\prod_{j=1}^{\infty} H_{k_j}(\langle x, e_j\rangle),$$

where $x \in B$, $k_j \neq 0$ for only finitely many j's, $k = \sum_{j=1}^{\infty} k_j$, H_{k_j} is the Hermite polynomial of degree k_j, and $\langle\cdot,\cdot\rangle = {}_B\langle\cdot,\cdot\rangle_{B^*}$ denotes the B-B^* duality pairing. We then denote by $\mathcal{H}_k$ the closure of the set of polynomial chaoses of order k under $L^2(B, \mu)$. The elements in $\mathcal{H}_k$ are called *homogeneous Wiener chaoses* of order k. We also set

$$\mathcal{H}_{\leq k} = \bigoplus_{j=0}^{k} \mathcal{H}_j \quad \text{for } k \in \mathbb{N}.$$

Let $L = \Delta - x\cdot\nabla$ be the Ornstein–Uhlenbeck operator.[20] Then it is known that any element in $\mathcal{H}_k$ is an eigenfunction of L with eigenvalue $-k$. As a consequence of the hypercontractivity of the Ornstein–Uhlenbeck semigroup $U(t) = e^{tL}$ due to Nelson [71], we have the following Wiener chaos estimate [92, Theorem I.22] (see also [94, Proposition 2.4]).

Lemma 2.5. *Let $k \in \mathbb{N}$. Then*

$$\|X\|_{L^p(\Omega)} \leq (p-1)^{k/2}\|X\|_{L^2(\Omega)} \quad \textit{for any } p \geq 2 \textit{ and any } X \in \mathcal{H}_{\leq k}.$$

The following lemma will be used in studying regularities of stochastic objects. We say that a stochastic process $X : \mathbb{R}_+ \to \mathcal{D}'(\mathbb{T}^d)$ is *spatially homogeneous* if $\{X(\cdot, t)\}_{t\in\mathbb{R}_+}$ and $\{X(x_0 + \cdot, t)\}_{t\in\mathbb{R}_+}$ have the same law for any $x_0 \in \mathbb{T}^d$. Given $h \in \mathbb{R}$, we define the difference operator δ_h by setting

$$\delta_h X(t) = X(t+h) - X(t). \tag{2.7}$$

[20] For simplicity, we write the definition of the Ornstein–Uhlenbeck operator L when $B = \mathbb{R}^d$.

Lemma 2.6. *Let $\{X_N\}_{N\in\mathbb{N}}$ and X be spatially homogeneous stochastic processes $\mathbb{R}_+ \to \mathcal{D}'(\mathbb{T}^d)$. Suppose that there exists $k \in \mathbb{N}$ such that $X_N(t)$ and $X(t)$ belong to $\mathcal{H}_{\leq k}$ for each $t \in \mathbb{R}_+$.*

(i) *Let $t \in \mathbb{R}_+$. If there exists $s_0 \in \mathbb{R}$ such that*
$$\mathbb{E}[|\widehat{X}(n,t)|^2] \lesssim \langle n\rangle^{-d-2s_0}$$
for any $n \in \mathbb{Z}^d$, then $X(t) \in W^{s,\infty}(\mathbb{T}^d)$ for all $s < s_0$ almost surely. Furthermore, if there exists $\gamma > 0$ such that
$$\mathbb{E}[|\widehat{X}_N(n,t) - \widehat{X}(n,t)|^2] \lesssim N^{-\gamma}\langle n\rangle^{-d-2s_0}$$
for any $n \in \mathbb{Z}^d$ and $N \geq 1$, then $X_N(t)$ converges to $X(t)$ in $W^{s,\infty}(\mathbb{T}^d)$ for all $s < s_0$ almost surely.

(ii) *Let $T > 0$ and suppose that* (i) *holds on* $[0,T]$. *If there exists $\sigma \in (0,1)$ such that*
$$\mathbb{E}[|\delta_h\widehat{X}(n,t)|^2] \lesssim \langle n\rangle^{-d-2s_0+\sigma}|h|^\sigma$$
for any $n \in \mathbb{Z}^d$, $t \in [0,T]$, and $h \in [-1,1]$,[21] *then $X \in C([0,T]; W^{s,\infty}(\mathbb{T}^d))$ for all $s < s_0 - \sigma/2$ almost surely. Furthermore, if there exists $\gamma > 0$ such that*
$$\mathbb{E}[|\delta_h\widehat{X}_N(n,t) - \delta_h\widehat{X}(n,t)|^2] \lesssim N^{-\gamma}\langle n\rangle^{-d-2s_0+\sigma}|h|^\sigma,$$
for any $n \in \mathbb{Z}^d$, $t \in [0,T]$, $h \in [-1,1]$, and $N \geq 1$, then X_N converges to X in $C([0,T]; W^{s,\infty}(\mathbb{T}^d))$ for all $s < s_0 - \sigma/2$ almost surely.

Lemma 2.6 follows from a straightforward application of the Wiener chaos estimate (Lemma 2.5). For the proof, see [69, Proposition 3.6] and [76, Appendix]. As compared with [69, Proposition 3.6], we made small adjustments. In studying the time regularity, we made the following modifications: $\langle n\rangle^{-d-2s_0+2\sigma} \mapsto \langle n\rangle^{-d-2s_0+\sigma}$ and $s < s_0 - \sigma \mapsto s < s_0 - \sigma/2$, which is suitable for studying the wave equation. Moreover, while the result in [69] is stated in terms of the Besov–Hölder space $\mathcal{C}^s(\mathbb{T}^d) = B^s_{\infty,\infty}(\mathbb{T}^d)$, Lemma 2.6 features the L^∞-based Sobolev space $W^{s,\infty}(\mathbb{T}^3)$. Note that the required modification of the proof is straightforward since $W^{s,\infty}(\mathbb{T}^d)$ and $B^s_{\infty,\infty}(\mathbb{T}^d)$ differ only logarithmically:
$$\|f\|_{W^{s,\infty}} \leq \sum_{j=0}^\infty \|\mathbf{P}_j f\|_{W^{s,\infty}} \lesssim \|f\|_{B^{s+\varepsilon}_{\infty,\infty}}$$
for any $\varepsilon > 0$. For the proof of the almost sure convergence claims, see [76].

Lastly, we recall the following theorem of Wick (see [92, Proposition I.2]).

Lemma 2.7. *Let $g_1, \dots, g_{2n}$ be* (*not necessarily distinct*) *real-valued jointly Gaussian random variables. Then*
$$\mathbb{E}[g_1 \cdots g_{2n}] = \sum \prod_{k=1}^n \mathbb{E}[g_{i_k} g_{j_k}],$$
where the sum is over all partitions of $\{1,\dots,2n\}$ into disjoint pairs (i_k, j_k).

[21] We impose $h \geq -t$, so that $t + h \geq 0$.

3. On the stochastic terms: Part I

In this and the next sections, we establish the regularity properties of the stochastic objects $\uparrow$, $\curlyvee$, and $\curlyvee\!\!\circ$ defined in (1.15), (1.16), and (1.27), respectively. The following lemma establishes the regularity properties of the stochastic convolution $\uparrow$ and the Wick power $\curlyvee$ (see also [39, proof of Proposition 2.1]).

Lemma 3.1. *Let $T > 0$.*

(i) *For any $\varepsilon > 0$, $\uparrow_N$ in (1.19) converges to $\uparrow$ in $C([0,T]; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely. In particular,*

$$\uparrow \in C([0,T]; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3)) \quad \textit{almost surely.}$$

(ii) *For any $\varepsilon > 0$, $\curlyvee_N$ in (1.20) converges to $\curlyvee$ in $C([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely. In particular,*

$$\curlyvee \in C([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3)) \quad \textit{almost surely.}$$

Proof. (i) Let $t \geq 0$. From (1.9), we have

$$\widehat{\uparrow}(n,t) = \int_0^t \frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle}\, d\beta_n(t') \tag{3.1}$$

and thus

$$\mathbb{E}[|\widehat{\uparrow}(n,t)|^2] = \sigma_n(t,t) = \frac{t}{2\langle n\rangle^2} - \frac{\sin(2t\langle n\rangle)}{4\langle n\rangle^3} \leq C(t)\langle n\rangle^{-2} \tag{3.2}$$

for any $n \in \mathbb{Z}^3$, where $\sigma_n(t,t)$ is defined in (1.33). Hence from Lemma 2.6, we conclude that $\uparrow(t) \in W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3)$ almost surely for any $\varepsilon > 0$.

Let $0 \leq t_2 \leq t_1$. From (1.9), we have

$$\begin{aligned}\widehat{\uparrow}(n,t_1) - \widehat{\uparrow}(n,t_2) &= \int_{t_2}^{t_1} \frac{\sin((t_1-t')\langle n\rangle)}{\langle n\rangle}\, d\beta_n(t') \\ &\quad + \int_0^{t_2} \frac{\sin((t_1-t')\langle n\rangle) - \sin((t_2-t')\langle n\rangle)}{\langle n\rangle}\, d\beta_n(t').\end{aligned} \tag{3.3}$$

Then, from the mean value theorem, we have

$$\begin{aligned}\mathbb{E}[|\widehat{\uparrow}(n,t_1) - \widehat{\uparrow}(n,t_2)|^2] &\lesssim \langle n\rangle^{-2}|t_1-t_2| + t_2\langle n\rangle^{-2+\sigma}|t_1-t_2|^\sigma \\ &\leq C(t_2)\langle n\rangle^{-2+\sigma}|t_1-t_2|^\sigma\end{aligned} \tag{3.4}$$

for any $n \in \mathbb{Z}^3$, $0 \leq t_2 \leq t_1$ with $t_1 - t_2 \leq 1$, and $\sigma \in [0,1]$. Hence, from Lemma 2.6, we conclude that $\uparrow \in C(\mathbb{R}_+; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely for any $\varepsilon > 0$.

Proceeding as above, we have

$$\mathbb{E}[|\widehat{\uparrow}_M(n,t) - \widehat{\uparrow}_N(n,t)|^2] \leq C(t)\mathbf{1}_{|n|>N}\cdot\langle n\rangle^{-2} \leq C(t)N^{-\gamma}\langle n\rangle^{-2+\gamma}.$$

for any $n \in \mathbb{Z}^3$, $M \geq N \geq 1$, and $\gamma \geq 0$. Similarly, with δ_h as in (2.7), we have

$$\begin{aligned}\mathbb{E}[|\delta_h\widehat{\uparrow}_M(n,t) - \delta_h\widehat{\uparrow}_N(n,t)|^2] &\lesssim C(t)\mathbf{1}_{|n|>N}\cdot\langle n\rangle^{-2+\sigma}|h|^\sigma \\ &\lesssim C(t)N^{-\gamma}\langle n\rangle^{-2+\sigma+\gamma}|h|^\sigma\end{aligned}$$

for any $n \in \mathbb{Z}^3$, $M \geq N \geq 1$, $h \in [-1,1]$, $\gamma \geq 0$, and $\sigma \in [0,1]$. Therefore, it follows from Lemma 2.6 that given $T > 0$ and $\varepsilon > 0$, the truncated stochastic convolution $\dot{\mid}_N$ converges to $\dot{\mid}$ in $C([0,T]; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely.

(ii) Proceeding as in part (i), the main task is to estimate $\mathbb{E}[|\widehat{\dot{\vee}}(n,t)|^2]$. The following discussion holds for $\dot{\vee}_N$ with constants independent of $N \in \mathbb{N} \cup \{\infty\}$. From (1.20) and (1.21), we have

$$\mathbb{E}[|\widehat{\dot{\vee}}(n,t)|^2] = \sum_{n_1+n_2=n} \sum_{n_1'+n_2'=n} \mathbb{E}\Big[\big(\widehat{\dot{\mid}}(n_1,t)\widehat{\dot{\mid}}(n_2,t) - \mathbf{1}_{n=0}\cdot \mathbb{E}[|\widehat{\dot{\mid}}(n_1,t)|^2]\big) \times \overline{\big(\widehat{\dot{\mid}}(n_1',t)\widehat{\dot{\mid}}(n_2',t) - \mathbf{1}_{n=0}\cdot \mathbb{E}[|\widehat{\dot{\mid}}(n_1',t)|^2]\big)}\Big]. \tag{3.5}$$

In order to have a nonzero contribution to (3.5), we must have $n_1 = n_1'$ and $n_2 = n_2'$ up to permutation. Thus, with (1.9) and Lemma 2.3, we have

$$\mathbb{E}[|\widehat{\dot{\vee}}(n,t)|^2] \lesssim t^2 \sum_{n=n_1+n_2} \frac{1}{\langle n_1\rangle^2 \langle n_2\rangle^2} \lesssim t^2 \langle n\rangle^{-1}. \tag{3.6}$$

Hence from Lemma 2.6, we conclude that $\dot{\vee}(t) \in W^{-1-\varepsilon,\infty}(\mathbb{T}^3)$ almost surely for any $\varepsilon > 0$. A similar argument shows that $\dot{\vee} \in C([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely and that $\dot{\vee}_N$ converges to $\dot{\vee}$ in $C([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely. ∎

Remark 3.2. As we saw in the proof of Lemma 3.1 (i), once we establish regularity properties of a given stochastic object τ, a slight modification of the argument shows convergence of the truncated stochastic objects τ_N to τ. Hence, in the following, we only establish claimed regularity properties of given stochastic terms.

Next, we study the regularity of $\dot{Y}$. As pointed out in the introduction, a naive parabolic thinking would give a regularity of $0- = (-\frac{1}{2}-) + (-\frac{1}{2}-) + 1$, where one degree of smoothing comes from the Duhamel integral operator $\mathcal{I}$. By exploiting the multilinear dispersive effect, we show that there is in fact an extra $\frac{1}{2}$-smoothing.

Proof of Proposition 1.6. By definition $\dot{Y} = \mathcal{I}(\dot{\vee})$, we have

$$\widehat{\dot{Y}}(n,t) = \int_0^t \frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle} \widehat{\dot{\vee}}(n,t')\,dt' \tag{3.7}$$

and thus

$$\widehat{\partial_t \dot{Y}}(n,t) = \int_0^t \cos((t-t')\langle n\rangle)\widehat{\dot{\vee}}(n,t')\,dt'.$$

Then, from (the proof of) Lemma 3.1 (ii), we conclude that

$$\partial_t \dot{Y} \in C([0,T]; W^{-1-\varepsilon,\infty}(\mathbb{T}^3)) \quad \text{almost surely for any } \varepsilon > 0.$$

In the following, we focus on proving that $\dot{Y} \in C([0,T]; W^{1/2-\varepsilon,\infty}(\mathbb{T}^3))$ almost surely. In view of Lemma 2.6, it suffices to show that there exists a small $\sigma \in (0,1)$

such that

$$\mathbb{E}[|\widehat{\curlyvee}(n,t)|^2] \le C(T)\langle n\rangle^{-4+}, \tag{3.8}$$

$$\mathbb{E}[|\widehat{\curlyvee}(n,t_1) - \widehat{\curlyvee}(n,t_2)|^2] \le C(T)\langle n\rangle^{-4+\sigma+}|t_1-t_2|^\sigma \tag{3.9}$$

for any $n \in \mathbb{Z}^3$ and $0 \le t, t_1, t_2 \le T$ with $0 < |t_1 - t_2| < 1$.

We first prove (3.8). From (3.7), we have

$$\mathbb{E}[|\widehat{\curlyvee}(n,t)|^2] = \int_0^t \frac{\sin((t-t_1)\langle n\rangle)}{\langle n\rangle}\int_0^t \frac{\sin((t-t_2)\langle n\rangle)}{\langle n\rangle}\mathbb{E}\big[\widehat{\vee}(n,t_1)\overline{\widehat{\vee}(n,t_2)}\big]\,dt_2\,dt_1. \tag{3.10}$$

When $n=0$, it follows from (1.20) with (1.21) and (1.33) that

$$\begin{aligned}\mathbb{E}[|\widehat{\curlyvee}(0,t)|^2] &= \int_0^t \sin(t-t_1)\int_0^t \sin(t-t_2)\mathbb{E}[\widehat{\vee}(0,t_1)\overline{\widehat{\vee}(0,t_2)}]\,dt_2\,dt_1\\ &= \int_0^t \sin(t-t_1)\int_0^t \sin(t-t_2)\\ &\quad\times \sum_{k_1,k_2\in\mathbb{Z}^3}\mathbb{E}\big[(|\widehat{\dagger}(k_1,t_1)|^2 - \sigma_{k_1}(t_1,t_1))(|\widehat{\dagger}(k_2,t_2)|^2 - \sigma_{k_2}(t_2,t_2))\big]\,dt_2\,dt_1\\ &\le C(T)\sum_{k\in\mathbb{Z}^3}\frac{1}{\langle k\rangle^4} \le C(T),\end{aligned}$$

where $\sigma_{k_j}(t_j,t_j)$ is as in (1.33). In the last step, we have used

$$\mathbb{E}\Big[\big(|\widehat{\dagger}(k_1,t_1)|^2 - \sigma_{k_1}(t_1,t_1)\big)\big(|\widehat{\dagger}(k_2,t_2)|^2 - \sigma_{k_2}(t_2,t_2)\big)\Big] = \mathbf{1}_{k_1=\pm k_2}\cdot\sigma_{k_1}(t_1,t_2)^2. \tag{3.11}$$

The identity (3.11) follows from Wick's theorem (Lemma 2.7). This proves (3.8) when $n=0$.

In the following, we assume $n\neq 0$. By expanding $\widehat{\vee}(n,t_1)$ and $\widehat{\vee}(n,t_2)$ as in (3.5) with $n = n_1 + n_2$ for $\widehat{\vee}(n,t_1)$ and $n = n_1' + n_2'$ for $\widehat{\vee}(n,t_2)$, we see that we must have $n_1 = n_1'$ and $n_2 = n_2'$ up to permutation in order to have a nonzero contribution to (3.10). Without loss of generality, assume that $0 \le t_2 \le t_1 \le t$. Then

$$\begin{aligned}&\mathbb{E}[|\widehat{\curlyvee}(n,t)|^2]\\ &\quad= 4\sum_{\substack{n_1+n_2=n\\ n_1\neq\pm n_2}}\int_0^t \frac{\sin((t-t_1)\langle n\rangle)}{\langle n\rangle}\int_0^{t_1}\frac{\sin((t-t_2)\langle n\rangle)}{\langle n\rangle}\sigma_{n_1}(t_1,t_2)\sigma_{n_2}(t_1,t_2)\,dt_2\,dt_1\\ &\qquad + 2\cdot\mathbf{1}_{n\in2\mathbb{Z}^3\setminus\{0\}}\int_0^t \frac{\sin((t-t_1)\langle n\rangle)}{\langle n\rangle}\int_0^{t_1}\frac{\sin((t-t_2)\langle n\rangle)}{\langle n\rangle}\\ &\qquad\qquad\times\mathbb{E}\big[\widehat{\dagger}(n/2,t_1)^2\overline{\widehat{\dagger}(n/2,t_2)^2}\big]\,dt_2\,dt_1\\ &\quad=: \mathrm{I}(n,t) + \mathrm{II}(n,t),\end{aligned} \tag{3.12}$$

where $\sigma_{n_j}(t_1,t_2)$ is as in (1.33) and $\mathrm{II}(n,t)$ denotes the contribution from $n_1 = n_2 = n_1' = n_2' = n/2$.

We first estimate the second term $\mathrm{II}(n,t)$ in (3.12). By Wick's theorem (Lemma 2.7) together with (1.33), we have

$$\big|\mathbb{E}\big[\widehat{\mathsf{1}}(n/2,t_1)^2\overline{\widehat{\mathsf{1}}(n/2,t_2)^2}\big]\big| \le C(T)\langle n\rangle^{-4}$$

for $0 \le t_2 \le t_1 \le t \le T$. Hence, from (3.12), we conclude that

$$|\mathrm{II}(n,t)| \le C(T)\langle n\rangle^{-6},$$

satisfying (3.8).

In the following, we estimate $\mathrm{I}(n,t)$ in (3.12):

$$\begin{aligned}\mathrm{I}(n,t) = -\sum_{k_1,k_2\in\{1,2\}}\sum_{\varepsilon_1,\varepsilon_2\in\{-1,1\}}\frac{\varepsilon_1\varepsilon_2 e^{i(\varepsilon_1+\varepsilon_2)t\langle n\rangle}}{\langle n\rangle^2}\sum_{\substack{n_1+n_2=n\\ n_1\ne\pm n_2}}\int_0^t e^{-i\varepsilon_1 t_1\langle n\rangle}\\ \times\int_0^{t_1} e^{-i\varepsilon_2 t_2\langle n\rangle}\prod_{j=1}^2\sigma_{n_j}^{(k_j)}(t_1,t_2)\,dt_2\,dt_1 =: \sum_{k_1,k_2\in\{1,2\}}\mathrm{I}^{(k_1,k_2)}(n,t),\end{aligned}\tag{3.13}$$

where

$$\begin{aligned}\sigma_n^{(1)}(t_1,t_2) &:= \frac{\cos((t_1-t_2)\langle n\rangle)}{2\langle n\rangle^2}t_2,\\ \sigma_n^{(2)}(t_1,t_2) &:= \frac{\sin((t_1-t_2)\langle n\rangle)}{4\langle n\rangle^3} - \frac{\sin((t_1+t_2)\langle n\rangle)}{4\langle n\rangle^3}\end{aligned}\tag{3.14}$$

so that

$$\sigma_n(t_1,t_2) = \sigma_n^{(1)}(t_1,t_2) + \sigma_n^{(2)}(t_1,t_2).$$

If $|n_1|\sim 1$ or $|n_2|\sim 1$, then from (3.14) with $\langle n_1\rangle\langle n_2\rangle \gtrsim \langle n\rangle$, we easily obtain

$$|\mathrm{I}(n,t)| \le C(T)\langle n\rangle^{-4+},\tag{3.15}$$

satisfying (3.8). Hence, we assume $|n_1|, |n_2| \gg 1$ in the following. By Lemma 2.3 with (3.14), we can easily bound the contribution to $\mathrm{I}(n,t)$ in (3.13) from $(k_1,k_2)\ne(1,1)$ and obtain for them the decay required in (3.15).

In the following, we estimate the worst contribution to $\mathrm{I}(n,t)$ coming from $(k_1,k_2) = (1,1)$:

$$\begin{aligned}\mathrm{I}^{(1,1)}(n,t) := -\frac{1}{16}\sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4\in\{-1,1\}}\sum_{\substack{n=n_1+n_2\\ n_1\ne\pm n_2}}\frac{\varepsilon_1\varepsilon_2 e^{i(\varepsilon_1+\varepsilon_2)t\langle n\rangle}}{\langle n\rangle^2\langle n_1\rangle^2\langle n_2\rangle^2}\\ \times\int_0^t e^{-it_1\kappa_1(\bar n)}\int_0^{t_1} t_2^2 e^{-it_2\kappa_2(\bar n)}\,dt_2\,dt_1,\end{aligned}$$

where

$$\begin{aligned}\kappa_1(\bar n) &:= \varepsilon_1\langle n\rangle - \varepsilon_3\langle n_1\rangle - \varepsilon_4\langle n_2\rangle,\\ \kappa_2(\bar n) &:= \varepsilon_2\langle n\rangle + \varepsilon_3\langle n_1\rangle + \varepsilon_4\langle n_2\rangle.\end{aligned}$$

When $|n| \lesssim 1$, (3.15) trivially holds. Hence, we assume $|n| \gg 1$. We have to carefully estimate the different contributions coming from the various combinations of $\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ by exploiting either (i) the dispersion (= oscillation) or (ii) smallness of the measure of the relevant frequency set.

Fix our choice of $\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ and denote by $\mathrm{I}^{(1,1)}_{\bar{\varepsilon}}(n,t)$ the associated contribution to $\mathrm{I}^{(1,1)}(n,t)$. By switching the order of integration and first integrating in t_1, we have

$$\begin{aligned}
&\bigg|\int_0^t e^{-it_1\kappa_1(\bar{n})} \int_0^{t_1} t_2^2 e^{-it_2\kappa_2(\bar{n})}\, dt_2\, dt_1\bigg| \\
&\qquad = \bigg|\int_0^t t_2^2 e^{-it_2\kappa_2(\bar{n})} \frac{e^{-it\kappa_1(\bar{n})} - e^{-it_2\kappa_1(\bar{n})}}{-i\kappa_1(\bar{n})}\, dt_2\bigg| \le C(T)(1+|\kappa_1(\bar{n})|)^{-1}.
\end{aligned}$$

Thus,

$$|\mathrm{I}^{(1,1)}_{\bar{\varepsilon}}(n,t)| \le C(T) \sum_{n_1+n_2=n} \frac{1}{\langle n\rangle^2\langle n_1\rangle^2\langle n_2\rangle^2(1+|\kappa_1(\bar{n})|)}. \tag{3.16}$$

Without loss of generality, by symmetry we can assume $|n_1| \ge |n_2|$ in the following when estimating the sum on the right-hand side.

Case 1: $(\varepsilon_1, \varepsilon_3, \varepsilon_4) = (\pm 1, \mp 1, \mp 1)$ or $(\pm 1, \mp 1, \pm 1)$. In this case, $|\kappa_1(\bar{n})| \ge \langle n\rangle$. Then, from Lemma 2.3, we obtain

$$|\mathrm{I}^{(1,1)}_{\bar{\varepsilon}}(n,t)| \le C(T)\langle n\rangle^{-4}.$$

This proves (3.8).

Case 2: $(\varepsilon_1, \varepsilon_3, \varepsilon_4) = (\pm 1, \pm 1, \mp 1)$. In this case, $|\kappa_1(\bar{n})| = \langle n\rangle + \langle n_2\rangle - \langle n_1\rangle$. For $n = n_1 + n_2$ and $|n_1| \ge |n_2|$, we have

$$\langle n_1\rangle \sim \langle n\rangle + \langle n_2\rangle. \tag{3.17}$$

When $n = n_1 + n_2$, the three vectors n, n_1, and n_2 form a triangle, where we view n_1 as a vector based at n_2. Then, by the law of cosines, we have

$$|n|^2 + |n_2|^2 - |n_1|^2 = 2|n|\,|n_2|\cos(\angle(n, n_2)). \tag{3.18}$$

From (3.17) and (3.18), we have

$$\begin{aligned}
|\kappa_1(\bar{n})| &= \frac{(\langle n\rangle + \langle n_2\rangle)^2 - \langle n_1\rangle^2}{\langle n\rangle + \langle n_2\rangle + \langle n_1\rangle} = \frac{2\langle n\rangle\langle n_2\rangle + |n|^2 + |n_2|^2 - |n_1|^2 + 1}{\langle n\rangle + \langle n_2\rangle + \langle n_1\rangle} \\
&\gtrsim \frac{|n|\,|n_2|(1-\cos\theta)}{\langle n_1\rangle}
\end{aligned} \tag{3.19}$$

where $\theta = \angle(n_2, -n) \in [0, \pi]$.

Subcase 2.i: $1 - \cos\theta \gtrsim 1$ (see Figure 1). In this case, from (3.16) and (3.19) with Lemma 2.3, we have

$$|\mathrm{I}_{\bar{\varepsilon}}^{(1,1)}(n,t)| \le C(T) \sum_{n_1+n_2=n} \frac{1}{\langle n\rangle^3 \langle n_1\rangle \langle n_2\rangle^3} \le C(T)\langle n\rangle^{-4+}, \tag{3.20}$$

yielding (3.8).

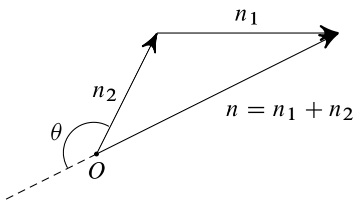

Fig. 1. A typical configuration in Subcase 2.i.

Subcase 2.ii: $1 - \cos\theta \ll 1$. In this case, we have $0 \le \theta \ll 1$, so n and n_2 point in almost opposite directions. In particular, we have $1 - \cos\theta \sim \theta^2 \ll 1$. By dyadically decomposing n_2 into $|n_2| \sim N_2$ for dyadic $N_2 \ge 1$, we see that for a fixed $n \in \mathbb{Z}^3$, the range of possible n_2 with $|n_2| \sim N_2$ is constrained to a cone $\mathcal{C}$ whose height is $\sim N_2 \cos\theta \sim N_2$ and the base disc of radius $\sim N_2 \sin\theta \sim N_2\theta$ with the direction of the central axis of the cone given by $-n$. Hence, we have $\mathrm{vol}(\mathcal{C}) \sim N_2^3\theta^2$ (see Figure 2). Then, from (3.16) and (3.19) with $|n_1| \gtrsim \max(|n|, |n_2|)$, we have

$$|\mathrm{I}_{\bar{\varepsilon}}^{(1,1)}(n,t)| \le C(T) \sum_{\substack{N_2 \ge 1\\ \text{dyadic}}} \frac{1}{\langle n\rangle^3 \max(\langle n\rangle, N_2) N_2^3\theta^2} N_2^3\theta^2 \le C(T)\langle n\rangle^{-4+}, \tag{3.21}$$

yielding (3.8).

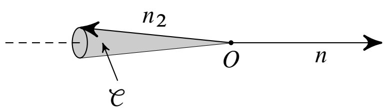

Fig. 2. A typical configuration in Subcase 2.ii. Here, we omit the vector n_1.

Case 3: $(\varepsilon_1, \varepsilon_3, \varepsilon_4) = (\pm 1, \pm 1, \pm 1)$. In this case, we have $|\kappa_1(\bar{n})| = \langle n_1\rangle + \langle n_2\rangle - \langle n\rangle$. By the law of cosines,

$$|n_1|^2 + |n_2|^2 - |n|^2 = 2|n_1|\,|n_2| \cos(\angle(-n_1, n_2)). \tag{3.22}$$

Then, by proceeding as in Case 2 with (3.17) and (3.22), we have

$$|\kappa_1(\bar{n})| = \frac{(\langle n_1\rangle + \langle n_2\rangle)^2 - \langle n\rangle^2}{\langle n_1\rangle + \langle n_2\rangle + \langle n\rangle} \gtrsim \frac{|n_1|\,|n_2|(1-\cos\theta)}{\langle n_1\rangle} \tag{3.23}$$

where $\theta = \angle(n_1, n_2) \in [0, \pi]$. When $1 - \cos\theta \gtrsim 1$, we can proceed as in (3.20). Next, consider the case $1 - \cos\theta \sim \theta^2 \ll 1$. Since $n = n_1 + n_2$, we see that $\angle(n, n_2) < \theta = \angle(n_1, n_2)$ in this case.[22] Hence, with $|n_1| \gtrsim |n|$, we can repeat the computation in (3.21) and obtain the same bound. This concludes the proof of (3.8) by choosing $\delta > 0$ sufficiently small.

Next, we briefly discuss the difference estimate (3.9). Let $0 \le t_2 \le t_1 \le T$. We need to estimate

$$\begin{aligned}\mathbb{E}[|\widehat{\text{Y}}(n,t_1) - \widehat{\text{Y}}(n,t_2)|^2] &= \mathbb{E}\big[(\widehat{\text{Y}}(n,t_1) - \widehat{\text{Y}}(n,t_2))\overline{\widehat{\text{Y}}(n,t_1)}\big] \\ &\quad - \mathbb{E}\big[(\widehat{\text{Y}}(n,t_1) - \widehat{\text{Y}}(n,t_2))\overline{\widehat{\text{Y}}(n,t_2)}\big].\end{aligned} \tag{3.24}$$

From (3.7), we have

$$\begin{aligned}\widehat{\text{Y}}(n,t_1) - \widehat{\text{Y}}(n,t_2) &= \int_{t_2}^{t_1} \frac{\sin((t_1 - t')\langle n\rangle)}{\langle n\rangle}\widehat{\text{V}}(n,t')\,dt' \\ &\quad + \int_0^{t_2} \frac{\sin((t_1 - t')\langle n\rangle) - \sin((t_2 - t')\langle n\rangle)}{\langle n\rangle}\widehat{\text{V}}(n,t')\,dt'.\end{aligned} \tag{3.25}$$

We crudely estimate (3.24) by using (3.25), (3.6), and the mean value theorem to control the difference. As a result, we have

$$\mathbb{E}[|\widehat{\text{Y}}(n,t_1) - \widehat{\text{Y}}(n,t_2)|^2] \lesssim C(T)\langle n\rangle^{-2}|t_1 - t_2|. \tag{3.26}$$

By interpolating (3.8) and (3.26), we obtain (3.9) for some small $\sigma \in (0, 1)$.

This completes the proof of Proposition 1.6. ∎

Remark 3.3. In Cases 2 and 3, we separately estimated the contributions from (i) $1 - \cos\theta \gtrsim 1$ and (ii) $1 - \cos\theta \ll 1$. Note that these correspond to the time non-resonant and (nearly) time resonant case in the dispersive PDE terminology. In the time resonant case (ii), there was no gain from time integration and thus we needed to exploit the smallness of the set (i.e. the cone $\mathcal{C}$) for the time resonant case by observing that the time resonance is caused by the parallel interaction of waves, i.e. n, n_1, and n_2 (close to) being parallel. A need for such a geometric consideration is one difference between the study of dispersive equations that of parabolic equations.

4. On the stochastic terms: Part II

In this section, we study the regularity property of the resonant product Y° defined in (1.27) (Proposition 1.8). From (1.23) and the definition of the Littlewood–Paley projector $\mathcal{F}(\mathbf{P}_j f)(n) = \varphi_j(n)\widehat{f}(n)$, we have

[22] Form a triangle with three vectors n, n_1, and n_2 with n and n_2 sharing a common base point such that $n = n_1 + n_2$. Then $\angle(n_1, n_2)$ is an exterior angle to this triangle and thus $\angle(n_1, n_2) = \angle(n, n_2) + \angle(-n, -n_1) > \angle(n, n_2)$.

$$\widehat{\mathbin{\text{Y}}_{\ominus}}(n,t) = \sum_{\substack{n_1+n_2+n_3=n\\ n_1+n_2\neq 0}} \sum_{|j-k|\leq 2} \varphi_j(n_1+n_2)\varphi_k(n_3) \times \int_0^t \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle} \widehat{\uparrow}(n_1,t')\widehat{\uparrow}(n_2,t')\,dt' \cdot \widehat{\uparrow}(n_3,t) + \sum_{n_1\in\mathbb{Z}^3}\sum_{|k|\leq 2}\varphi_k(n_3)\int_0^t \sin(t-t')\cdot\big(|\widehat{\uparrow}(n_1,t')|^2-\sigma_{n_1}(t')\big)\,dt'\cdot\widehat{\uparrow}(n,t) =: \widehat{\mathcal{R}}_1(n,t)+\widehat{\mathcal{R}}_2(n,t).$$

For simplicity of notation, however, we write

$$\widehat{\mathcal{R}}_1(n,t) = \sum_{\substack{n_1+n_2+n_3=n\\ |n_1+n_2|\sim|n_3|\\ n_1+n_2\neq 0}} \int_0^t \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle}\widehat{\uparrow}(n_1,t')\widehat{\uparrow}(n_2,t')\,dt'\cdot\widehat{\uparrow}(n_3,t),$$
$$\widehat{\mathcal{R}}_2(n,t) = \sum_{n_1\in\mathbb{Z}^3}\mathbf{1}_{|n|\sim 1}\int_0^t \sin(t-t')\cdot\big(|\widehat{\uparrow}(n_1,t')|^2-\sigma_{n_1}(t')\big)dt'\cdot\widehat{\uparrow}(n,t), \tag{4.1}$$

where the conditions $|n_1+n_2|\sim|n_3|$ in the first term and $|n|\sim 1$ in the second term signify the resonant product $\ominus$. The second term $\mathcal{R}_2$ in (4.1) corresponds to the contribution from $n_1+n_2=0$ and is already renormalized by the Wick renormalization: $\uparrow^2 \rightsquigarrow \mathbin{\text{V}}$. Using (3.11) and Lemma 2.6, it is easy to see that $\mathcal{R}_2 \in C(\mathbb{R}_+;C^\infty(\mathbb{T}^3))$ almost surely, since $|n|\sim 1$.

In the following, our main goal is to show

$$\mathbb{E}[|\widehat{\mathcal{R}}_1(n,t)|^2] \leq C(t)\langle n\rangle^{-3+}. \tag{4.2}$$

Then Lemma 2.6 allows us to conclude that $\mathcal{R}_1(t)\in W^{0-,\infty}(\mathbb{T}^3)$ almost surely. We decompose $\mathcal{R}_1$ as

$$\begin{aligned}\widehat{\mathcal{R}}_1(n,t) &= \sum_{\substack{n_1+n_2+n_3=n\\ |n_1+n_2|\sim|n_3|\\ (n_1+n_2)(n_2+n_3)(n_3+n_1)\neq 0}} \int_0^t \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle}\widehat{\uparrow}(n_1,t')\widehat{\uparrow}(n_2,t')\,dt'\cdot\widehat{\uparrow}(n_3,t)\\ &\quad + 2\int_0^t \widehat{\uparrow}(n,t')\Bigg[\sum_{\substack{n_2\in\mathbb{Z}^3\\ |n_2|\sim|n+n_2|\neq 0}}\frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\times\big(\widehat{\uparrow}(n_2,t')\widehat{\uparrow}(-n_2,t)-\sigma_{n_2}(t,t')\big)\Bigg]dt'\\ &\quad + 2\int_0^t \widehat{\uparrow}(n,t')\Bigg[\sum_{\substack{n_2\in\mathbb{Z}^3\\ |n_2|\sim|n+n_2|\neq 0}}\frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\sigma_{n_2}(t,t')\Bigg]dt'\\ &\quad - \mathbf{1}_{n\neq 0}\int_0^t \frac{\sin((t-t')\langle 2n\rangle)}{\langle 2n\rangle}(\widehat{\uparrow}(n,t'))^2dt'\cdot\widehat{\uparrow}(-n,t)\\ &=: \widehat{\mathcal{R}}_{11}(n,t)+\widehat{\mathcal{R}}_{12}(n,t)+\widehat{\mathcal{R}}_{13}(n,t)+\widehat{\mathcal{R}}_{14}(n,t).\end{aligned} \tag{4.3}$$

In this sum, the second term $\mathcal{R}_{12}$ corresponds to the "renormalized" contribution from $n_1 + n_3 = 0$ or $n_2 + n_3 = 0$, while the fourth term corresponds to the contribution from $n_1 = n_2 = n = -n_3$.

From (3.1), we have

$$\mathbb{E}[|\widehat{\mathcal{R}}_{14}(n,t)|^2] \le C(t)\langle n\rangle^{-8},$$

satisfying (4.2). Under $|n + n_2| \sim |n_2|$, we have $|n_2| \gtrsim |n|$. Then, using a variant of (3.11), we obtain

$$\mathbb{E}[|\widehat{\mathcal{R}}_{12}(n,t)|^2] \le C(t) \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2|}} \frac{1}{\langle n\rangle^2 \langle n_2\rangle^6} \lesssim \langle n\rangle^{-5},$$

satisfying (4.2).

Given $n \in \mathbb{Z}^3$, define

$$\begin{aligned} NR(n) := \{(n_1, n_2, n_3) \in \mathbb{Z}^3 : n = n_1 + n_2 + n_3,\ |n_1 + n_2| \sim |n_3|, \\ (n_1 + n_2)(n_2 + n_3)(n_3 + n_1) \ne 0\}. \end{aligned}$$

Then, with shorthand notation $n_{ij} = n_i + n_j$, we have

$$\begin{aligned} &\mathbb{E}[|\widehat{\mathcal{R}}_{11}(n,t)|^2] \\ &= \mathbb{E}\bigg[\sum_{(n_1,n_2,n_3)\in NR(n)} \int_0^t \frac{\sin((t-t_1)\langle n_{12}\rangle)}{\langle n_{12}\rangle} \widehat{\uparrow}(n_1,t_1)\widehat{\uparrow}(n_2,t_1)\, dt' \cdot \widehat{\uparrow}(n_3,t) \\ &\quad \times \sum_{(n_1',n_2',n_3')\in NR(n)} \int_0^t \frac{\sin((t-t_2)\langle n_{12}'\rangle)}{\langle n_{12}'\rangle} \overline{\widehat{\uparrow}(n_1',t_2)\widehat{\uparrow}(n_2',t_2)}\, dt' \cdot \overline{\widehat{\uparrow}(n_3',t)}\bigg]. \end{aligned}$$

In order to compute the expectation above, we need to take all possible pairings between (n_1, n_2, n_3) and (n_1', n_2', n_3'). By Jensen's inequality, however, we see that it suffices to consider the case $n_j = n_j'$, $j = 1, 2, 3$ (see the discussion on $\curlyvee$ in [69, Section 4]; see also [45, Section 10]). Hence, by Wick's theorem and (3.10), we have

$$\mathbb{E}[|\widehat{\mathcal{R}}_{11}(n,t)|^2] \lesssim \sum_{\substack{m+n_3=n \\ |m|\sim|n_3|}} \mathbb{E}[|\widehat{\curlyvee}(m,t)|^2]\mathbb{E}[|\widehat{\uparrow}(n_3,t)|^2].$$

From (3.2), (3.8), and Lemma 2.3 (ii), we have

$$\begin{aligned} \mathbb{E}[|\widehat{\mathcal{R}}_{11}(n,t)|^2] &\le C(t) \sum_{\substack{m+n_3=n \\ |m|\sim|n_3|}} \frac{1}{\langle m\rangle^{4-}\langle n_3\rangle^2} \\ &\le C(t)\langle n\rangle^{-3+}, \end{aligned}$$

proving (4.2). Note that in evaluating the last sum, we have crucially used the fact that the product is a resonant product.

It remains to study the third term $\widehat{\mathcal{R}}_{13}$ on the right-hand side of (4.3). Let $0 \le t_2 \le t_1 \le T$. Then, from (4.3) and (1.33), we have

$$\begin{aligned}\mathbb{E}[|\widehat{\mathcal{R}}_{13}(n,t)|^2] = 8 &\sum_{k_0,k_1,k_2\in\{1,2\}} \int_0^t \int_0^{t_1} \sigma_n^{(k_0)}(t_1,t_2)\\ &\times \Bigg[\sum_{\substack{n_2\in\mathbb{Z}^3\\ |n_2|\sim|n+n_2|\neq 0}} \frac{\sin((t-t_1)\langle n+n_2\rangle)}{\langle n+n_2\rangle}\sigma_{n_2}^{(k_1)}(t,t_1)\Bigg]\\ &\times \Bigg[\sum_{\substack{n_2'\in\mathbb{Z}^3\\ |n_2'|\sim|n+n_2'|\neq 0}} \frac{\sin((t-t_2)\langle n+n_2'\rangle)}{\langle n+n_2'\rangle}\sigma_{n_2'}^{(k_2)}(t,t_2)\Bigg] dt_2\, dt_1\\ =: &\sum_{k_0,k_1,k_2\in\{1,2\}} \mathrm{I}^{(k_0,k_1,k_2)}(n,t),\end{aligned}$$

where $\sigma_n(t,t') = \sigma_n^{(1)}(t,t') + \sigma_n^{(2)}(t,t')$ as in (3.14). In the following, we only consider the contribution from $(k_0,k_1,k_2) = (1,1,1)$ since the other cases follow in a similar (but easier) manner.

From (3.14), we have

$$\begin{aligned}\mathrm{I}^{(1,1,1)}(n,t) = &\int_0^t \int_0^{t_1} \frac{\cos((t_1-t_2)\langle n\rangle)}{\langle n\rangle^2} t_2\\ &\times \Bigg[\sum_{\substack{n_2\in\mathbb{Z}^3\\ |n_2|\sim|n+n_2|\neq 0}} \frac{\sin((t-t_1)\langle n+n_2\rangle)}{\langle n+n_2\rangle} \frac{\cos((t-t_1)\langle n_2\rangle)}{\langle n_2\rangle^2} t_1\Bigg]\\ &\times \Bigg[\sum_{\substack{n_2'\in\mathbb{Z}^3\\ |n_2'|\sim|n+n_2'|\neq 0}} \frac{\sin((t-t_2)\langle n+n_2'\rangle)}{\langle n+n_2'\rangle} \frac{\cos((t-t_2)\langle n_2'\rangle)}{\langle n_2'\rangle^2} t_2\Bigg] dt_2\, dt_1\\ = -\frac{1}{32} &\sum_{\substack{\varepsilon_j\in\{-1,1\}\\ j=1,\dots,5}} \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n_2|\sim|n+n_2|\neq 0}} \sum_{\substack{n_2'\in\mathbb{Z}^3\\ |n_2'|\sim|n+n_2'|\neq 0}} \frac{\varepsilon_1\varepsilon_2 e^{it(\varepsilon_1\langle n+n_2\rangle+\varepsilon_2\langle n+n_2'\rangle+\varepsilon_3\langle n_2\rangle+\varepsilon_4\langle n_2'\rangle)}}{\langle n\rangle^2\langle n+n_2\rangle\langle n_2\rangle^2\langle n+n_2'\rangle\langle n_2'\rangle^2}\\ &\times \int_0^t t_1 e^{-it_1\kappa_3(\bar n)} \int_0^{t_1} t_2^2 e^{-it_2\kappa_4(\bar n')} dt_2\, dt_1,\end{aligned} \tag{4.4}$$

where

$$\begin{aligned}\kappa_3(\bar n) &:= \varepsilon_1\langle n+n_2\rangle + \varepsilon_3\langle n_2\rangle - \varepsilon_5\langle n\rangle,\\ \kappa_4(\bar n') &:= \varepsilon_2\langle n+n_2'\rangle + \varepsilon_4\langle n_2'\rangle + \varepsilon_5\langle n\rangle.\end{aligned} \tag{4.5}$$

Under the constraint $|n+n_2| \sim |n_2|$ and $|n+n_2'| \sim |n_2'|$, we have $|n_2|, |n_2'| \gtrsim |n|$. In the following, we also assume $|n_2| \gtrsim |n_2'|$.

We decompose $\mathrm{I}^{(1,1,1)}(n,t)$ according to the value of $\bar{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_5) \in \{\pm 1\}^5$ and write

$$\mathrm{I}^{(1,1,1)}(n,t) =: \sum_{\bar{\varepsilon} \in \{\pm 1\}^5} \mathrm{I}^{(1,1,1)}_{\bar{\varepsilon}}(n,t).$$

In the following, we study $\mathrm{I}^{(1,1,1)}_{\bar{\varepsilon}}$ for each fixed $\bar{\varepsilon} \in \{\pm 1\}^5$. Note that at first glance, the sums over n_2 and n_2' in (4.4) do not seem to be absolutely convergent. In many cases, we make use of dispersion (i.e. time oscillation) and show that they are indeed absolutely convergent. In Case 3 below, however, there is a subcase where we show that the sum is only conditionally convergent. In this case, it is understood that the sum is first studied under the constraint $|n_2|, |n_2'| \le N$ for some $N \ge 1$ and that the sum remains bounded in taking a limit $N \to \infty$. We do not mention this procedure in an explicit manner in the following.

By first integrating (4.4) in t_1 when $|\kappa_3(\bar{n})| \ge 1$ and simply bounding the integral in (4.4) by $C(T)$ when $|\kappa_3(\bar{n})| < 1$, we have

$$|\mathrm{I}^{(1,1,1)}_{\bar{\varepsilon}}(n,t)| \le \frac{C(T)}{\langle n \rangle^2} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2|}} \frac{1}{\langle n+n_2 \rangle \langle n_2 \rangle^2 (1 + |\kappa_3(\bar{n})|)} \times \sum_{\substack{n_2' \in \mathbb{Z}^3 \\ |n+n_2'| \sim |n_2'|}} \frac{\mathbf{1}_{\{|n_2| \gtrsim |n_2'|\}}}{\langle n+n_2' \rangle \langle n_2' \rangle^2}. \tag{4.6}$$

Case 1: $(\varepsilon_1, \varepsilon_3, \varepsilon_5) = (\pm 1, \pm 1, \mp 1)$. In this case, (4.5) implies that $|\kappa_3(\bar{n})| \gtrsim \langle n_2 \rangle$. By writing $(1 + |\kappa_3(\bar{n})|)^{-1} \lesssim \langle n \rangle^{-1+2\delta} \langle n_2 \rangle^{-\delta} \langle n_2' \rangle^{-\delta}$ in (4.6) for sufficiently small $\delta > 0$ and applying Lemma 2.3, we obtain

$$|\mathrm{I}^{(1,1,1)}_{\bar{\varepsilon}}(n,t)| \le C(T) \langle n \rangle^{-3}. \tag{4.7}$$

Case 2: $(\varepsilon_1, \varepsilon_3, \varepsilon_5) = (\pm 1, \pm 1, \pm 1)$. If $|n+n_2| \sim |n_2| \gg |n|$, then $|\kappa_3(\bar{n})| \gtrsim \langle n_2 \rangle$ and hence (4.7) holds as above. Otherwise, $|n+n_2| \sim |n_2| \sim |n|$. In this case, $\langle n \rangle^{-2\delta} \lesssim \langle n_2 \rangle^{-\delta} \langle n_2' \rangle^{-\delta}$ for $\delta > 0$. Then

$$\begin{aligned} |\mathrm{I}^{(1,1,1)}_{\bar{\varepsilon}}(n,t)| &\le \frac{C(T)}{\langle n \rangle^{2-2\delta}} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2|}} \frac{1}{\langle n+n_2 \rangle \langle n_2 \rangle^{2+\delta} (1 + |\kappa_3(\bar{n})|)} \times \sum_{\substack{n_2' \in \mathbb{Z}^3 \\ |n+n_2'| \sim |n_2'|}} \frac{\mathbf{1}_{\{|n_2| \gtrsim |n_2'|\}}}{\langle n+n_2' \rangle \langle n_2' \rangle^{2+\delta}} \\ &\le \frac{C(T)}{\langle n \rangle^{2-\delta}} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2|}} \frac{1}{\langle n+n_2 \rangle \langle n_2 \rangle^{2+\delta} (1 + |\kappa_3(\bar{n})|)}. \end{aligned} \tag{4.8}$$

We can now proceed as in Case 3 of the proof of Proposition 1.6 by replacing (n, n_1, n_2) with $(n, n+n_2, -n_2)$. In particular, from (3.23), we have

$$|\kappa_3(\bar{n})| \gtrsim |n_2|(1 - \cos\theta) \tag{4.9}$$

where $\theta = \angle(n + n_2, -n_2) \in [0, \pi]$. When $1 - \cos\theta \gtrsim 1$, by summing over n_2 in (4.8) with (4.9) and Lemma 2.3, we obtain (4.7).

Next, consider the case $1 - \cos\theta \sim \theta^2 \ll 1$. Since $n = (n + n_2) + (-n_2)$, we see that the angle $\theta_0 = \angle(n, -n_2)$ is smaller than $\theta = \angle(n + n_2, -n_2)$ in this case. Moreover, we see that for fixed $n \in \mathbb{Z}^3$, the range of possible $-n_2$ with $|n_2| \sim N_2$, $N_2 \geq 1$ dyadic, is constrained to a cone whose height is $\sim N_2|\cos\theta_0| \sim N_2$ and the base disc has radius $\sim N_2 \sin\theta_0 \lesssim N_2\theta$. Then, from (4.8) and (4.9) with $|n + n_2| \sim |n_2| \sim |n|$, we have

$$|\mathrm{I}_{\varepsilon}^{(1,1,1)}(n,t)| \leq \frac{C(T)}{\langle n\rangle^{2-\delta}} \sum_{\substack{N_2 \sim \langle n\rangle \\ \text{dyadic}}} \frac{1}{N_2^{4+\delta}\,\theta^2} N_2^3\theta^2 \leq C(T)\langle n\rangle^{-3+\delta},$$

yielding (4.7).

Case 3: $\varepsilon_1 = -\varepsilon_3$. First, suppose that $|n| \geq |n_2|^\gamma$ for some small $\gamma > 0$ (to be chosen later). Then, with $\langle n\rangle^{-2\delta} \lesssim \langle n_2\rangle^{-\gamma\delta}\langle n_2'\rangle^{-\gamma\delta}$ for $\lambda > 0$, we can proceed as in Case 2 (but using the computation of Case 2 in the proof of Proposition 1.6 by replacing (n, n_1, n_2) with $(n, n + n_2, -n_2)$ or $(n, -n_2, n + n_2)$) and obtain

$$\begin{aligned}|\mathrm{I}_{\bar\varepsilon}^{(1,1,1)}(n,t)| \leq \frac{C(T)}{\langle n\rangle^{2-2\delta}} &\sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2|}} \frac{1}{\langle n + n_2\rangle\langle n_2\rangle^{2+\gamma\delta}(1 + |\kappa_3(\bar n)|)} \\ &\times \sum_{\substack{n_2' \in \mathbb{Z}^3 \\ |n+n_2'| \sim |n_2'|}} \frac{\mathbf{1}_{\{|n_2| \gtrsim |n_2'|\}}}{\langle n + n_2'\rangle\langle n_2'\rangle^{2+\gamma\delta}} \\ \leq C(T)\langle n\rangle^{-3+(2-\gamma)\delta},\end{aligned}$$

sayisfying (4.2) by choosing $\delta > 0$ sufficiently small.

Next, we consider the case $|n| \ll |n_2|^\gamma$. In this case, we are not able to prove absolute summability in (4.6) since $\kappa_3(\bar n)$ does not have any good lower bound, and thus we need to proceed more carefully. By writing out the contribution from the sum over n_2 in (4.4) (namely, ignoring the sum over n_2'), we have

$$\int_0^t t_1 e^{it_1\varepsilon_5\langle n\rangle} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n+n_2| \sim |n_2| \\ |n| \ll |n_2|^\gamma}} \frac{\sin((t - t_1)(\langle n + n_2\rangle - \langle n_2\rangle))}{\langle n\rangle^2\langle n + n_2\rangle\langle n_2\rangle^2}\, dt_1. \tag{4.10}$$

By going back to the definition (1.23) of the resonant product $\circleddash$, we can write down the sum over n_2 in (4.10) as

$$\sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n| \ll |n_2|^\gamma}} \varphi_j(n_2) \sum_{|k-j| \leq 2} \varphi_k(n + n_2), \tag{4.11}$$

where φ_j is as in (2.1). Thanks to the restriction $|n| \ll |n_2|^\gamma$ with small $\gamma > 0$, the sum in (4.11) is in fact given by

$$\sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n| \ll |n_2|^\gamma}} \varphi_j(n_2). \tag{4.12}$$

While the sum over n_2 in (4.10) is not absolutely convergent, we do not expect to gain anything from the time integration in t_1 in this case due to the lack of a good lower bound on $\kappa_3(\bar{n})$. The reduction to (4.12), however, allows us to exploit the symmetry $n_2 \leftrightarrow -n_2$ and the oscillatory nature of the sine kernel in (4.10).

By the Taylor remainder theorem, we have

$$\Theta^{\pm}(n, n_2) := \langle n \pm n_2 \rangle - \langle n_2 \rangle \mp \frac{\langle n, n_2 \rangle}{\langle n_2 \rangle} = O\bigg(\frac{\langle n \rangle^2}{\langle n_2 \rangle} \bigg), \tag{4.13}$$

where $\langle n, n_2 \rangle = \langle n, n_2 \rangle_{\mathbb{R}^3}$ denotes the standard inner product on $\mathbb{R}^3$. Let Λ be the index set "$\sim \mathbb{Z}^3/2$" in (1.7). Then, with (4.13) and the mean value theorem, we have

$$\begin{aligned}
(4.10) &= \int_0^t t_1 e^{it_1 \varepsilon_5 \langle n \rangle} \sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \mathbb{Z}^3 \\ |n| \ll |n_2|^\gamma}} \varphi_j(n_2) \frac{\sin((t-t_1)(\langle n+n_2 \rangle - \langle n_2 \rangle))}{\langle n \rangle^2 \langle n+n_2 \rangle \langle n_2 \rangle^2} dt_1 \\
&= \int_0^t t_1 e^{it_1 \varepsilon_5 \langle n \rangle} \sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \Lambda \\ |n| \ll |n_2|^\gamma}} \frac{\varphi_j(n_2)}{\langle n \rangle^2 \langle n+n_2 \rangle \langle n_2 \rangle^2} \\
&\qquad \times \big\{ \sin((t-t_1)(\langle n+n_2 \rangle - \langle n_2 \rangle)) + \sin((t-t_1)(\langle n-n_2 \rangle - \langle n_2 \rangle)) \big\} dt_1 \\
&= \int_0^t t_1 e^{it_1 \varepsilon_5 \langle n \rangle} \sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \Lambda \\ |n| \ll |n_2|^\gamma}} \frac{\varphi_j(n_2)}{\langle n \rangle^2 \langle n+n_2 \rangle \langle n_2 \rangle^2} \\
&\qquad \times \bigg\{ \sin\bigg((t-t_1) \bigg(\frac{\langle n, n_2 \rangle}{\langle n_2 \rangle} + \Theta^+(n, n_2) \bigg) \bigg) \\
&\qquad\qquad - \sin\bigg((t-t_1) \bigg(\frac{\langle n, n_2 \rangle}{\langle n_2 \rangle} - \Theta^-(n, n_2) \bigg) \bigg) \bigg\} dt_1 \\
&\le C(T) \sum_{j \in \mathbb{N}_0} \sum_{\substack{n_2 \in \Lambda \\ |n| \ll |n_2|^\gamma}} \frac{\varphi_j(n_2)}{\langle n \rangle^2 \langle n+n_2 \rangle \langle n_2 \rangle^2} \frac{\langle n \rangle^{4\delta}}{\langle n_2 \rangle^{2\delta}} \\
&\le \frac{C(T)}{\langle n_2' \rangle^{\delta}} \sum_{\substack{n_2 \in \Lambda \\ |n| \ll |n_2|^\gamma}} \frac{1}{\langle n \rangle^{2-4\delta} \langle n_2 \rangle^{3+\delta}}
\end{aligned} \tag{4.14}$$

for any $\delta \in [0, 1/2]$. Fix a small $\delta > 0$. By applying Lemma 2.3 to sum over n_2 and n_2' and then using the condition $|n| \ll |n_2|^\gamma$, we obtain

$$|\mathrm{I}_{\bar{\varepsilon}}^{(1,1,1)}(n, t)| \le \frac{C(T)}{\langle n \rangle^{2-3\delta}} \cdot \frac{1}{\langle n \rangle^{\delta/\gamma}} \le C(T) \langle n \rangle^{-3}$$

by choosing $\gamma = \gamma(\delta) > 0$ sufficiently small. This proves (4.2).

Next, we briefly discuss the difference estimate. In view of Lemma 2.6, we need to show that there exists $\sigma \in (0, 1)$ such that

$$\mathbb{E}[|\widehat{\mathcal{R}}_1(n, t_1) - \widehat{\mathcal{R}}_1(n, t_2)|^2] \le C(T) \langle n \rangle^{-3+\sigma+} |t_1 - t_2|^{\sigma} \tag{4.15}$$

for any $n \in \mathbb{Z}^3$ and $0 \le t, t_1, t_2 \le T$ with $0 < |t_1 - t_2| < 1$. As in (3.24), we need to estimate

$$\mathbb{E}[|\widehat{\mathcal{R}}_1(n,t_1) - \widehat{\mathcal{R}}_1(n,t_2)|^2] \\ = \sum_{j=1}^{2} (-1)^{j+1} \mathbb{E}\big[(\widehat{\mathcal{R}}_1(n,t_1) - \widehat{\mathcal{R}}_1(n,t_2)) \overline{\widehat{\Upsilon}(n,t_j)}\big]. \tag{4.16}$$

As for $\mathcal{R}_{11}$, $\mathcal{R}_{12}$, and $\mathcal{R}_{14}$ in (4.3), we can crudely estimate them and obtain

$$\mathbb{E}[|\widehat{\mathcal{R}}_{1j}(n,t_1) - \widehat{\mathcal{R}}_{1j}(n,t_2)|^2] \lesssim C(T)\langle n\rangle^{-2+}|t_1 - t_2| \tag{4.17}$$

for $j = 1, 2, 4$, since the relevant summations are absolutely convergent. Then (4.15) follows from interpolating (4.17) and

$$\mathbb{E}[|\widehat{\mathcal{R}}_{1j}(n,t_1) - \widehat{\mathcal{R}}_{1j}(n,t_2)|^2] \lesssim C(T)\langle n\rangle^{-3+}.$$

It remains to discuss $\mathcal{R}_{13}$. In view of (4.4) and (4.16), we need to consider an expression like

$$\int_0^{t_2}\int_0^{\tau_1} \frac{\cos((\tau_1-\tau_2)\langle n\rangle)}{\langle n\rangle^2} \tau_2 \Bigg[\sum_{\substack{n_2\in\mathbb{Z}^3 \\ |n+n_2|\sim|n_2|}} \frac{\sin((t-\tau_1)\langle n+n_2\rangle)}{\langle n+n_2\rangle} \frac{\cos((t-\tau_1)\langle n_2\rangle)}{\langle n_2\rangle^2} \tau_1 \Bigg]\Bigg|_{t=t_2}^{t_1} \\ \times \Bigg[\sum_{\substack{n_2'\in\mathbb{Z}^3 \\ |n+n_2'|\sim|n_2'|}} \frac{\sin((t_j-\tau_2)\langle n+n_2'\rangle)}{\langle n+n_2'\rangle} \frac{\cos((t_j-\tau_2)\langle n_2'\rangle)}{\langle n_2'\rangle^2} \tau_2 \Bigg] d\tau_2\, d\tau_1$$

for $0 \le \tau_2 \le \tau_1 \le T$. Then, by repeating the computations in Cases 1–3 above and applying the mean value theorem, we directly obtain (4.15). Note that some of the relevant summations are not absolutely convergent in this case and hence we cannot proceed with a crude estimate and interpolation. Compare this with the parabolic setting (see [69, Section 5]).

This completes the proof of Proposition 1.8.

5. Paracontrolled operators

We first present the proof of Lemma 1.9 on the regularity of $Z = (S(t)(X_0, X_1)) \circledcirc \uparrow$.

Proof of Lemma 1.9. Let

$$H(t) = S(t)(X_0, X_1).$$

For $n = n_1 + n_2$ and $|n_1| \sim |n_2|$, we have $\langle n\rangle \lesssim \langle n_1\rangle \sim \langle n_2\rangle$. Then it follows from Minkowski's integral inequality, the Wiener chaos estimate (Lemma 2.5), independence of $\widehat{\uparrow}(n_2, t)$, and (3.2) that for any $p \ge 2$, we have

$$\begin{aligned}
\| \|Z(t)\|_{H^s} \|_{L^p(\Omega)} &\le \Big\| \| \langle n\rangle^s \mathcal{F}_x(H \circleddash \Upsilon)(n,t)\|_{L^p(\Omega)} \Big\|_{\ell^2_n} \\
&\le p^{1/2} \Big\| \| \langle n\rangle^s \mathcal{F}_x(H \circleddash \Upsilon)(n,t)\|_{L^2(\Omega)} \Big\|_{\ell^2_n} \\
&\sim p^{1/2} \Big(\sum_{n\in\mathbb{Z}^3} \langle n\rangle^{2s} \sum_{\substack{n_1+n_2=n\\ |n_1|\sim|n_2|}} |\widehat{H}(n_1,t)|^2 \mathbb{E}[|\widehat{\Upsilon}(n_2,t)|^2] \Big)^{1/2} \\
&\lesssim p^{1/2} \Big(\sum_{n\in\mathbb{Z}^3} \langle n\rangle^{2(s-s_1-1)} \sum_{n_1\in\mathbb{Z}^3} \langle n_1\rangle^{2s_1} |\widehat{H}(n_1,t)|^2 \Big)^{1/2} \\
&\lesssim p^{1/2} \|(X_0,X_1)\|_{\mathcal{H}^{s_1}}
\end{aligned} \tag{5.1}$$

provided that $s < s_1 - 1/2$. Fix $\varepsilon > 0$ small. Then, by writing

$$(H \circleddash \Upsilon)(t_1) - (H \circleddash \Upsilon)(t_2) = H(t_1) \circleddash (\Upsilon(t_1) - \Upsilon(t_2)) + (H(t_1) - H(t_2)) \circleddash \Upsilon(t_2)$$

for $0 \le t_2 \le t_1$, we can repeat the computation in (5.1). In particular, by (3.4) and the mean value theorem, we obtain

$$\begin{aligned}
&\Big\| \|(H \circleddash \Upsilon)(t_1) - (H \circleddash \Upsilon)(t_2)\|_{H^s} \Big\|_{L^p(\Omega)} \\
&\qquad \lesssim C(t_2) p^{1/2} |t_1 - t_2|^{\varepsilon/2} \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + p^{1/2} \|H(t_1) - H(t_2)\|_{H^{s_1-\varepsilon/2}} \\
&\qquad \lesssim C(t_2) p^{1/2} |t_1 - t_2|^{\varepsilon/2} \|(X_0,X_1)\|_{\mathcal{H}^{s_1}}
\end{aligned}$$

provided that $s < s_1 - 1/2 - \varepsilon/2$. Therefore, by taking large $p = p(\varepsilon) \gg 1$, we conclude from Kolmogorov's continuity criterion [6, Theorem 8.2] that Z belongs to $C([0,T]; H^{s_1-1/2-\varepsilon}(\mathbb{T}^3))$ almost surely. ∎

The remaining part of this section is devoted to the mapping properties of the paracontrolled operators $\mathfrak{I}^{(1)}_{\circleddash}$ in (1.30) and $\mathfrak{I}_{\circledless,\circleddash}$ in (1.31).

We first study the regularity properties of $\mathfrak{I}^{(1)}_{\circledless}$. By writing out the frequency relation $|n_2|^\theta \lesssim |n_1| \ll |n_2|$ more carefully, we have

$$\begin{aligned}
\mathfrak{I}^{(1)}_{\circledless}(w)(t) = \sum_{n\in\mathbb{Z}^3} e_n \sum_{n_1+n_2=n} \sum_{\theta k + c_0 \le j < k-2} &\varphi_j(n_1)\varphi_k(n_2) \\
&\times \int_0^t \frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle} \widehat{w}(n_1,t') \widehat{\Upsilon}(n_2,t')\, dt',
\end{aligned} \tag{5.2}$$

where $c_0 \in \mathbb{R}$ is some fixed constant. In the following, we establish the mapping property of $\mathfrak{I}^{(1)}_{\circledless}$ in a deterministic manner by using pathwise regularity of the stochastic convolution Υ.

Given $\Xi \in C(\mathbb{R}_+; W^{-1/2-\varepsilon}(\mathbb{T}^3))$ for some small $\varepsilon > 0$, define a paracontrolled operator $\mathfrak{I}^{(1),\Xi}_{\circledless}$ by

$$\begin{aligned}
\mathfrak{I}^{(1),\Xi}_{\circledless}(w)(t) := \sum_{n\in\mathbb{Z}^3} e_n \sum_{n_1+n_2=n} \sum_{\theta k + c_0 \le j < k-2} &\varphi_j(n_1)\varphi_k(n_2) \\
&\times \int_0^t \frac{\sin((t-t')\langle n\rangle)}{\langle n\rangle} \widehat{w}(n_1,t') \widehat{\Xi}(n_2,t')\, dt'.
\end{aligned} \tag{5.3}$$

Note that we have $\mathfrak{I}^{(1)}_{\circledless} = \mathfrak{I}^{(1),\Upsilon}_{\circledless}$, i.e. with $\Xi = \Upsilon$.

Lemma 5.1. *Let $0 < s_1 < 1/2$ and $T > 0$. Then, given a small $\theta > 0$, there exists a small $\varepsilon = \varepsilon(s_1, \theta) > 0$ such that given $\Xi \in C(\mathbb{R}_+; W^{-1/2-\varepsilon,\infty}(\mathbb{T}^3))$, the paracontrolled operator $\mathfrak{I}_{\otimes}^{(1),\Xi}$ defined in (5.3) belongs to the class*

$$\mathcal{L}_2 := \mathcal{L}(C([0,T]; H^{s_1}(\mathbb{T}^3))\,;\, C([0,T]; H^{1/2+2\varepsilon}(\mathbb{T}^3)). \tag{5.4}$$

As a direct consequence of Lemmas 5.1 and 3.1, we obtain the following corollary for the paracontrolled operator $\mathfrak{I}_{\otimes}^{(1)}$ defined in (1.30) (and (5.2)).

Corollary 5.2. *Let $0 < s_1 < 1/2$ and $T > 0$. Then, given a small $\theta > 0$, there exists a small $\varepsilon = \varepsilon(s_1, \theta) > 0$ such that the paracontrolled operator $\mathfrak{I}_{\otimes}^{(1)}$ defined in (1.30) belongs to $\mathcal{L}_2$ of (5.4) almost surely. Moreover, by letting $\mathfrak{I}_{\otimes}^{(1),N}$, $N \in \mathbb{N}$, denote the paracontrolled operator in (1.30) with $\uparrow$ replaced by the truncated stochastic convolution $\uparrow_N$ of (1.19), the truncated paracontrolled operator $\mathfrak{I}_{\otimes}^{(1),N}$ converges almost surely to $\mathfrak{I}_{\otimes}^{(1)}$ in $\mathcal{L}_2$.*

Proof of Lemma 5.1. Let $s_1 > 0$. For $|n_2|^\theta \lesssim |n_1| \ll |n_2|$ with $n = n_1 + n_2$, we have

$$\langle n\rangle^{1/2+2\varepsilon}\frac{1}{\langle n\rangle} \lesssim \langle n_1\rangle^{4\varepsilon/\theta}\langle n_2\rangle^{-1/2-2\varepsilon} \lesssim \langle n_1\rangle^{s_1-\varepsilon}\langle n_2\rangle^{-1/2-2\varepsilon} \tag{5.5}$$

by choosing $\varepsilon = \varepsilon(s_1, \theta) > 0$ sufficiently small.

Let $\widehat{w}_j(n_1, t') = \varphi_j(n_1)\widehat{w}(n_1, t')$ and $\widehat{\Xi}_k(n_2, t') = \varphi_k(n_2)\widehat{\Xi}(n_2, t')$. Then, from (5.3) and (5.5), we have

$$\begin{aligned}
&\|\mathfrak{I}_{\otimes}^{(1),\Xi}(w)(t)\|_{H^{1/2+2\varepsilon}} \\
&\qquad \lesssim \int_0^t \sum_{j,k=0}^\infty 2^{(s_1-\varepsilon)j}2^{(-1/2-2\varepsilon)k}\bigg\|\sum_{n=n_1+n_2}\widehat{w}_j(n_1,t')\widehat{\Xi}_k(n_2,t')\bigg\|_{\ell^2_n} dt' \\
&\qquad \lesssim \int_0^t \sum_{j,k=0}^\infty 2^{(s_1-\varepsilon)j}2^{(-1/2-2\varepsilon)k}\|w_j(t')\|_{L^2_x}\|\Xi_k(t')\|_{L^\infty_x} dt'.
\end{aligned}$$

Then, by summing over dyadic blocks and applying the trivial embedding (2.5), we obtain

$$\begin{aligned}
\|\mathfrak{I}_{\otimes}^{(1),\Xi}(w)(t)\|_{H^{1/2+2\varepsilon}} &\lesssim T\|w\|_{L^\infty_T H^{s_1}_x}\|\Xi\|_{L^\infty_T(B^{-1/2-2\varepsilon}_{\infty,1})_x} \\
&\lesssim T\|w\|_{L^\infty_T H^{s_1}_x}\|\Xi\|_{L^\infty_T W^{-1/2-\varepsilon,\infty}_x}
\end{aligned}$$

for any $t \in [0, T]$. The continuity in time of $\mathfrak{I}_{\otimes}^{(1),\Xi}(w)$ follows from modifying the computation above as in the previous subsections. We omit the details. ∎

Finally, we present the proof of Proposition 1.11 on the paracontrolled operator $\mathfrak{I}_{\otimes,\ominus}$ in (1.31). By writing out the frequency relations more carefully as in (5.2), we have

$$\mathfrak{I}_{\otimes,\ominus}(w)(t) = \sum_{n\in\mathbb{Z}^3} e_n \int_0^t \sum_{j=0}^\infty \sum_{n_1\in\mathbb{Z}^3} \varphi_j(n_1)\widehat{w}(n_1,t')\mathcal{A}_{n,n_1}(t,t')\,dt', \tag{5.6}$$

where

$$\mathcal{A}_{n,n_1}(t,t') := \mathbf{1}_{[0,t]}(t') \sum_{\substack{k=0\\ j\le\theta k+c_0}}^{\infty} \sum_{\substack{\ell,m=0\\ |\ell-m|\le 2}}^{\infty} \sum_{n_2+n_3=n-n_1} \varphi_k(n_2)\varphi_\ell(n_1+n_2)\varphi_m(n_3) \times \frac{\sin((t-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle} \widehat{\mathfrak{1}}(n_2,t')\widehat{\mathfrak{1}}(n_3,t). \tag{5.7}$$

For ease of notation, however, we simply use (1.31) and (1.32) in the following, with the understanding that the frequency relations $|n_1| \ll |n_2|^\theta$ and $|n_1+n_2| \sim |n_3|$ are indeed characterized by the use of smooth frequency cutoff functions as in (5.6) and (5.7). Moreover, we drop the cutoff function $\mathbf{1}_{[0,t]}(t')$ in the following with the understanding that $0 \le t' \le t$.

Proof of Proposition 1.11. We separately consider the contributions to $\mathfrak{I}_{\circledless,\circleddash}$ from $\mathcal{A}^{(1)}_{n,n_1}$ and $\mathcal{A}^{(2)}_{n,n_1}$ defined in (1.34) and denote them respectively by $\mathfrak{I}^{(1)}_{\circledless,\circleddash}(w)$ and $\mathfrak{I}^{(2)}_{\circledless,\circleddash}(w)$.

Given a dyadic $N_2 \ge 1$, let $\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')$ be the contribution to $\mathcal{A}^{(1)}_{n,n_1}(t,t')$ from $\{|n_2| \sim N_2\}$.[23] Fix $0 \le t \le T$. Then, from (1.31) and (1.34), we have

$$\begin{aligned}
\|\mathfrak{I}^{(1)}_{\circledless,\circleddash}(w)(t)\|_{H^{s_2-1}} &\le \bigg\| \int_0^t \langle n\rangle^{s_2-1} \sum_{n_1\in\mathbb{Z}^3} \widehat{w}(n_1,t')\mathcal{A}^{(1)}_{n,n_1}(t,t')\,dt' \bigg\|_{\ell^2_n} \\
&\lesssim T^{1/2}\|w\|_{L^\infty_t L^2_x} \big\|\langle n\rangle^{s_2-1}\mathcal{A}^{(1)}_{n,n_1}(t,t')\big\|_{L^2_{t'}([0,T];\ell^2_{n,n_1})} \\
&\lesssim T^{1/2}\|w\|_{L^\infty_t L^2_x} \sum_{\substack{N_2\ge 1\\ \text{dyadic}}} \big\|\langle n\rangle^{s_2-1}\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')\big\|_{L^2_{t'}([0,T];\ell^2_{n,n_1})} \\
&\lesssim T^{1/2}\|w\|_{L^\infty_t L^2_x} \|\mathcal{A}^{(1)}(t,\cdot)\|_{\mathcal{A}(T)},
\end{aligned}$$

where we have introduced the norm

$$\|\mathcal{A}^{(1)}(t,\cdot)\|_{\mathcal{A}(T)} := \bigg(\sum_{\substack{N_2\ge 1\\ \text{dyadic}}} N_2^\delta \big\|\langle n\rangle^{s_2-1}\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')\big\|^2_{L^2_{t'}([0,T];\ell^2_{n,n_1})} \bigg)^{1/2}. \tag{5.8}$$

Remark 5.3. For fixed $t,t' \in [0,T]$, set

$$\mathcal{T}_{t,t'}(f) = \sum_{n_1\in\mathbb{Z}^3} \widehat{f}(n_1)\mathcal{A}^{(1)}_{n,n_1}(t,t')e_n.$$

Then the expression $\|\langle n\rangle^{s_2-1}\mathcal{A}^{(1)}_{n,n_1}(t,t')\|_{\ell^2_{n,n_1}}$ is none other than the Hilbert–Schmidt norm of the operator $\mathcal{T}_{t,t'}$ from $L^2(\mathbb{T}^3)$ into $H^{s_2-1}(\mathbb{T}^3)$. Recalling that the Hilbert–Schmidt norm of a given operator controls its operator norm, it is natural to work with the $\mathcal{A}(T)$-norm of $\mathcal{A}^{(1)}(t,\cdot)$ defined above (which is conveniently modified to carry out analysis on each dyadic block $\{|n_2| \sim N_2\}$).

[23]More precisely, $\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')$ denotes the contribution to (5.7) from $2^k \sim N_2$.

By a similar argument, we obtain

$$\begin{aligned}\|\mathfrak{I}^{(1)}_{\circledless,\circleddash}(w)(t_1) - \mathfrak{I}^{(1)}_{\circledless,\circleddash}(w)(t_2)\|_{H^{s_2-1}} &\lesssim |t_1 - t_2|^{1/2}\|w\|_{L^\infty_t L^2_x}\|\mathcal{A}^{(1)}(t_1,\cdot)\|_{\mathcal{A}(T)} \\ &\quad + T^{1/2}\|w\|_{L^\infty_t L^2_x}\|\mathcal{A}^{(1)}(t_1,\cdot) - \mathcal{A}^{(1)}(s_2,\cdot)\|_{\mathcal{A}(T)} \qquad (5.9)\end{aligned}$$

for $t_1, t_2 \in [0, T]$.

We now show that the random process $t \mapsto \mathcal{A}^{(1)}(t,\cdot)$ has almost surely continuous trajectories (in t) with respect to the Banach space generated by the norm $\|\cdot\|_{\mathcal{A}(T)}$. In order to do so, we apply Kolmogorov's continuity criterion and we need, as usual, to evaluate sufficiently high moments of the random variable $\mathcal{A}^{(1)}(t_1,\cdot) - \mathcal{A}^{(1)}(t_2,\cdot)$ with $t_1, t_2 \in [0, T]$.

Note that the conditions $|n_1| \ll |n_2|^\theta$ for some small $\theta > 0$ and $|n_1 + n_2| \sim |n_3|$ imply $|n_2| \sim |n_3| \gg |n_1|$. Moreover, since $n - n_1 = n_2 + n_3$, we have $|n_2| \sim |n_3| \gtrsim |n|$. Then, by (1.33), we have

$$\begin{aligned}&\mathbb{E}[\|\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')\|^2_{L^2_{t'}([0,T])}] \\ &\le \Bigg\| \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|\\ |n_2|\sim N_2\\ n_2+n_3\neq 0}} \frac{|\sin((t-t')\langle n_1+n_2\rangle)|^2}{\langle n_1+n_2\rangle^2}\mathbb{E}[|\widehat{\mathsf{1}}(n_2,t')\,\widehat{\mathsf{1}}(n_3,t)|^2]\Bigg\|_{L^1_{t'}([0,T])} \\ &\quad + \Bigg\| \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta\\ |n_2|\sim N_2}} \frac{|\sin((t-t')\langle n+n_2\rangle)|^2}{\langle n+n_2\rangle^2}\mathbb{E}[|\widehat{\mathsf{1}}(n_2,t')\,\widehat{\mathsf{1}}(-n_2,t) - \sigma_{n_2}(t,t')|^2]\Bigg\|_{L^1_{t'}([0,T])} \\ &\lesssim T\Bigg\{ \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|\\ |n_2|\sim N_2\\ n_1\neq n}} \frac{1}{\langle n_2\rangle^6} + \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta\\ |n_2|\sim N_2}} \frac{1}{\langle n+n_2\rangle^2\langle n_2\rangle^4}\Bigg\} \lesssim TN_2^{-3}\mathbf{1}_{|n_1|\ll N_2^\theta}\mathbf{1}_{|n|\lesssim N_2}. \qquad (5.10)\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathbb{E}[\|\mathcal{A}^{(1)}(t,\cdot)\|^2_{\mathcal{A}(T)}] &= \sum_{\substack{N_2\ge 1\\ \text{dyadic}}} N_2^\delta \sum_{n,n_1}\langle n\rangle^{2s_2-2}\mathbb{E}\big[\|\mathcal{A}^{(1)}_{n,n_1,N_2}(t,t')\|^2_{L^2_{t'}([0,T])}\big] \\ &\le \sum_{\substack{N_2\ge 1\\ \text{dyadic}}} N_2^{\delta-3} \sum_{n,n_1\in\mathbb{Z}^3}\langle n\rangle^{2s_2-2}\,\mathbf{1}_{|n_1|\ll N_2^\theta}\mathbf{1}_{|n|\lesssim N_2} \\ &\le \sum_{\substack{N_2\ge 1\\ \text{dyadic}}} N_2^{\delta+3\theta+2s_2-2} < \infty \qquad (5.11)\end{aligned}$$

by choosing $\delta = \delta(s_2) > 0$ and $\theta = \theta(s_2) > 0$ sufficiently small since $s_2 < 1$.

From (1.34), we have

$$\begin{aligned}
&\mathcal{A}^{(1)}_{n,n_1}(t_1,t') - \mathcal{A}^{(1)}_{n,n_1}(t_2,t')\\
&\quad = \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|}} \frac{\sin((t_1-t')\langle n_1+n_2\rangle) - \sin((t_2-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle} B_{n_2,n_3}(t_1,t')\\
&\qquad + \sum_{\substack{n_2+n_3=n-n_1\\ |n_1|\ll|n_2|^\theta\\ |n_1+n_2|\sim|n_3|}} \frac{\sin((t_2-t')\langle n_1+n_2\rangle)}{\langle n_1+n_2\rangle}\big(B_{n_2,n_3}(t_1,t') - B_{n_2,n_3}(t_2,t')\big),
\end{aligned} \tag{5.12}$$

where $B_{n_2,n_3}(t,t') = \widehat{\uparrow}(n_2,t')\,\widehat{\uparrow}(n_3,t) - \mathbf{1}_{n_2+n_3=0}\cdot\sigma_{n_2}(t,t')$. Arguing as in (5.10) and (5.11), we obtain

$$\mathbb{E}[\|\mathcal{A}^{(1)}(t_1,\cdot) - \mathcal{A}^{(1)}(t_2,\cdot)\|^2_{\mathcal{A}(T)}] \lesssim |t_1-t_2|^\sigma$$

for some small $\sigma > 0$. Indeed, the first term on the right-hand side of (5.12) can be controlled by the mean value theorem, creating an additional factor of $\langle n_1+n_2\rangle^\sigma|t_1-t_2|^\sigma$. On the other hand, by writing

$$\begin{aligned}
B_{n_2,n_3}(t_1,t') - B_{n_2,n_3}(t_2,t') &= \widehat{\uparrow}(n_2,t')(\widehat{\uparrow}(n_3,t_1) - \widehat{\uparrow}(n_3,t_2))\\
&\quad - \mathbb{E}\big[\widehat{\uparrow}(n_2,t')\,(\widehat{\uparrow}(-n_2,t_1) - \widehat{\uparrow}(-n_2,t_2))\big]\\
&\quad - \mathbf{1}_{n_2+n_3=0}\{\sigma_{n_2}(t_1,t') - \sigma_{n_2}(t_2,t')\},
\end{aligned}$$

we can apply (3.3) and the mean value theorem to create $|t_1-t_2|^\sigma$ at the expense of losing a small power in n_2 or n_3.

Finally, note that $\mathcal{A}^{(1)}_{n,n_1}$ is a homogeneous Wiener chaos of order 2. Therefore by Minkowski's inequality and the Wiener chaos estimate (Lemma 2.5), we obtain

$$\mathbb{E}[\|\mathcal{A}^{(1)}(t_1,\cdot) - \mathcal{A}^{(1)}(t_2,\cdot)\|^p_{\mathcal{A}(T)}] \lesssim p^p|t_1-t_2|^{\sigma p}$$

for any $p \geq 2$. Finally, by Kolmogorov's continuity criterion, we conclude that

$$\|\mathcal{A}^{(1)}(t_1,\cdot) - \mathcal{A}^{(1)}(t_2,\cdot)\|_{\mathcal{A}(T)} \leq C(\omega)|t_1-t_2|^{\sigma-} \tag{5.13}$$

for all $t_1,t_2 \in [0,T]$, where the constant $C = C(\omega)$ lies in $L^p(\Omega)$ for some large $p \gg 1$. From (5.9) and (5.13), we then deduce the required almost sure continuity for $\mathfrak{I}^{(1)}_{\circledcirc,\ominus}(w)$.

Next, we consider the contribution from $\mathcal{A}^{(2)}_{n,n_1}(t,t')$ in (1.34). This part is entirely deterministic. Since $\mathcal{A}^{(2)}_{n,n_1}(t,t') = 0$ unless $n = n_1$, we only consider $\mathcal{A}^{(2)}_{n,n}(t,t')$. In view of (3.14), we decompose $\mathcal{A}^{(2)}_{n,n}(t,t')$ as

$$\begin{aligned}\mathcal{A}^{(2)}_{n,n}(t,t') &= t' \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\frac{\cos((t-t')\langle n_2\rangle)}{2\langle n_2\rangle^2}\\ &\quad + \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\cdot O\bigg(\frac{1}{\langle n_2\rangle^3}\bigg).\\ &= t' \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')(\langle n+n_2\rangle+\langle n_2\rangle))}{4\langle n+n_2\rangle\langle n_2\rangle^2}\\ &\quad + t' \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')(\langle n+n_2\rangle-\langle n_2\rangle))}{4\langle n+n_2\rangle\langle n_2\rangle^2}\\ &\quad + \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')\langle n+n_2\rangle)}{\langle n+n_2\rangle}\cdot O\bigg(\frac{1}{\langle n_2\rangle^3}\bigg)\\ &=: \mathcal{A}^{(3)}_n(t,t') + \mathcal{A}^{(4)}_n(t,t') + \mathcal{A}^{(5)}_n(t,t'). \end{aligned} \tag{5.14}$$

We will show that

$$\|\langle n\rangle^{-\varepsilon}\mathcal{A}^{(j)}_n(t,t')\|_{\ell^\infty_n(\mathbb{Z}^3)} \le C(T) \tag{5.15}$$

for any $\varepsilon>0$, $0\le t'\le t\le T$, and $j=4,5$. Then, by denoting by $\mathfrak{I}^{(j)}_{\circledless,\circleddash}(w)$ the contribution to $\mathfrak{I}_{\circledless,\circleddash}(w)$ from $\mathbf{1}_{n=n_1}\cdot\mathcal{A}^{(j)}_n$, it follows from (1.31) and (5.15) that

$$\begin{aligned}\|\mathfrak{I}^{(j)}_{\circledless,\circleddash}(w)(t)\|_{H^{s_2-1}} &\lesssim \bigg\|\int_0^t \frac{1}{\langle n\rangle^{1-s_2}}\widehat{w}(n,t')\mathcal{A}^{(j)}_n(t,t')\,dt'\bigg\|_{\ell^2_n}\\ &\lesssim T\|w\|_{L^\infty_t L^2_x}\bigg\|\frac{1}{\langle n\rangle^{1-s_2}}\mathcal{A}^{(j)}_n(t,t')\bigg\|_{L^\infty_{t,t'}([0,T];\ell^\infty_n)}\\ &\lesssim C(T)\|w\|_{L^\infty_t L^2_x}\end{aligned} \tag{5.16}$$

for $t\in[0,T]$ and $j=4,5$, provided that $s_2<1$. The continuity in time of $\mathfrak{I}^{(j)}_{\circledless,\circleddash}(w)(t)$ follows from a similar argument.

By noting that $\langle n+n_2\rangle\sim\langle n_2\rangle\gg\langle n\rangle$ if $|n|\ll|n_2|^\theta$, we easily see that (5.15) is satisfied for $j=5$. On the other hand, the sum for $\mathcal{A}^{(4)}_n(t,t')$ is not absolutely convergent. As in Case 3 in Section 4, we exploit the symmetry $n_2\leftrightarrow -n_2$ and the oscillatory nature of the sine kernel. With (4.13) and the mean value theorem, we have

$$
\begin{aligned}
\mathcal{A}_n^{(4)}(t,t') &= t' \sum_{\substack{n_2\in\Lambda\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')(\langle n+n_2\rangle - \langle n_2\rangle)) + \sin((t-t')(\langle n-n_2\rangle - \langle n_2\rangle))}{4\langle n+n_2\rangle\langle n_2\rangle^2}\\
&\quad - \sum_{\substack{n_2\in\Lambda\\ |n|\ll|n_2|^\theta}} \frac{\sin((t-t')(\langle n-n_2\rangle - \langle n_2\rangle))}{4\langle n_2\rangle^2}\bigg(\frac{1}{\langle n+n_2\rangle} - \frac{1}{\langle n-n_2\rangle}\bigg)\\
&= \sum_{\substack{n_2\in\Lambda\\ |n|\ll|n_2|^\theta}} \frac{1}{4\langle n+n_2\rangle\langle n_2\rangle^2}\bigg\{\sin\bigg((t-t')\bigg(\frac{\langle n,n_2\rangle}{\langle n_2\rangle} + \Theta^+(n,n_2)\bigg)\bigg)\\
&\qquad - \sin\bigg((t-t')\bigg(\frac{\langle n,n_2\rangle}{\langle n_2\rangle} - \Theta^-(n,n_2)\bigg)\bigg)\bigg\} + O\bigg(\sum_{\substack{n_2\in\Lambda\\ |n|\ll|n_2|^\theta}} \frac{\langle n\rangle}{\langle n_2\rangle^4}\bigg)\\
&\lesssim \sum_{\substack{n_2\in\Lambda\\ |n|\ll|n_2|^\theta}} \frac{1}{\langle n+n_2\rangle\langle n_2\rangle^2}\frac{\langle n\rangle^{2\delta}}{\langle n_2\rangle^\delta} + O(1) \lesssim \langle n\rangle^\delta
\end{aligned}
\tag{5.17}
$$

for any $\delta \in (0,1]$. This proves (5.15) and hence (5.16) for $j=4$.

It remains to consider $\mathcal{A}_n^{(3)}(t,t')$. For this term, there is no internal cancellation structure and we need to make use of its fast oscillation by directly studying $\mathfrak{I}^{(3)}_{\circledcirc,\ominus}(w)$. From (1.31) and (5.14), we have

$$
\begin{aligned}
\|\mathfrak{I}^{(3)}_{\circledcirc,\ominus}(w)(t)\|_{H^{s_2-1}} \lesssim \sum_{\varepsilon_1\in\{-1,1\}} \bigg\| &\frac{e^{i\varepsilon_1 t(\langle n+n_2\rangle + \langle n_2\rangle)}}{\langle n\rangle^{1-s_2}}\\
&\times \sum_{\substack{n_2\in\mathbb{Z}^3\\ |n|\ll|n_2|^\theta}} \int_0^t \widehat{w}(n,t')t'\frac{e^{-i\varepsilon_1 t'(\langle n+n_2\rangle+\langle n_2\rangle)}}{\langle n+n_2\rangle\langle n_2\rangle^2}dt'\bigg\|_{\ell^2_n}.
\end{aligned}
\tag{5.18}
$$

Integrating by parts, we have

$$
\begin{aligned}
\int_0^t &\widehat{w}(n,t')t'\frac{e^{-i\varepsilon_1 t'(\langle n+n_2\rangle+\langle n_2\rangle)}}{\langle n+n_2\rangle\langle n_2\rangle^2}dt'\\
&= \frac{1}{-i\varepsilon_1(\langle n+n_2\rangle + \langle n_2\rangle)\langle n+n_2\rangle\langle n_2\rangle^2}\\
&\quad\times\bigg\{\widehat{w}(n,t)te^{-i\varepsilon_1 t(\langle n+n_2\rangle+\langle n_2\rangle)}\\
&\qquad - \int_0^t (\widehat{w}(n,t') + t'\partial_t\widehat{w}(n,t'))e^{-i\varepsilon_1 t'(\langle n+n_2\rangle+\langle n_2\rangle)}dt'\bigg\}.
\end{aligned}
\tag{5.19}
$$

Hence, from (5.18) and (5.19), we obtain

$$
\begin{aligned}
\|\mathfrak{I}^{(3)}_{\circledcirc,\ominus}(w)(t)\|_{H^{s_2-1}} &\lesssim \sum_{n_2\in\mathbb{Z}^3} \frac{1}{\langle n_2\rangle^{4-(s_2+\varepsilon)\theta}}(\|w\|_{L^\infty_T H^{-1-\varepsilon}_x} + \|\partial_t w\|_{L^\infty_T H^{-1-\varepsilon}_x})\\
&\lesssim \|w\|_{L^\infty_T H^{-1-\varepsilon}_x} + \|\partial_t w\|_{L^\infty_T H^{-1-\varepsilon}_x}
\end{aligned}
$$

for some small $\varepsilon > 0$. The continuity in time of $\mathfrak{I}^{(3)}_{\circledless,\circleddash}(w)(t)$ follows from a similar argument.

This completes the proof of Proposition 1.11. ■

Remark 5.4. In handling the term $\mathcal{A}^{(4)}_n(t,t')$ in (5.17), we exploited the symmetrization $n_2 \leftrightarrow -n_2$. This may seem to raise a possible issue in treating regularization via mollification when a mollification kernel does not satisfy certain symmetry properties (contrary to our claim in Remark 1.14). We point out, however, that this is not the case.

Recall that $\mathcal{A}^{(4)}_n(t,t')$ is defined in (5.14) from $\mathcal{A}^{(2)}_{n,n_1}(t,t')$, which in turn is defined in (1.34). Given a mollification kernel ρ, define ρ_δ as in (1.44). Then, defining the smoothed stochastic convolution $\uparrow_\delta = \mathcal{I}(\xi_\delta)$ with $\xi_\delta = \rho_\delta * \xi$, we construct $\mathcal{A}^{(2)}_{n,n_1}(t,t')$ associated with $\uparrow_\delta$. In this case, instead of (1.33), we have

$$\sigma^\delta_{n_2}(t_1,t_2) := \mathbb{E}[\widehat{\uparrow}_\delta(n_2,t_1)\widehat{\uparrow}_\delta(-n_2,t_2)] = \mathbb{E}[\widehat{\rho}_\delta(n_2)\widehat{\uparrow}(n_2,t_1)\widehat{\rho}_\delta(-n_2)\widehat{\uparrow}(-n_2,t_2)].$$

Note that the effect of mollification $\widehat{\eta}_\delta(n_2)\widehat{\eta}_\delta(-n_2)$ is symmetric in $n_2 \leftrightarrow -n_2$. This observation allows us to carry out the symmetrization argument in (5.17) (and also the symmetrization argument in (4.14) for the construction of $\curlyvee$) even for general (nonsymmetric) mollification.

Remark 5.5. We can also accommodate a space-time regularization of the noise in the form of a smoothing kernel $\rho(x,t)$. In this case, we need to impose an additional assumption that the space-time mollification kernel $\rho(x,t)$ is even in x, $\rho(-x,t) = \rho(x,t)$ for any $t \in \mathbb{R}$, implying that

$$\widehat{\rho}(-n,t) = \widehat{\rho}(n,t) \tag{5.20}$$

for any $n \in \mathbb{Z}^3$ and $t \in \mathbb{R}$.

This is not directly apparent from the computations above, so let us give some indications of the argument. In order to treat a space-time mollification, it is more convenient to switch to a representation of the stochastic objects based directly on the white noise. We write the stochastic convolution $\uparrow_\delta = \mathcal{I}(\xi_\delta) = \mathcal{I}(\rho_\delta * \xi)$ as

$$\widehat{\uparrow}_\delta(n,t) := \int_{\mathbb{R}}\bigg[\int_0^t \sum_{\varepsilon\in\{-1,1\}} \frac{\varepsilon}{2i}\, \frac{e^{i\varepsilon\langle n\rangle(t-t')}}{\langle n\rangle}\widehat{\rho}_\delta(n,t'-\tau)\,dt'\bigg]\,d\beta_n(\tau), \tag{5.21}$$

where $\widehat{\rho}_\delta(n,t)$ is the spatial Fourier transform of the space-time mollifier ρ_δ:

$$\rho_\delta(x,t) = \delta^{-4}\rho(\delta^{-1}x,\delta^{-1}t).$$

The renormalizations and computations of the various stochastic objects then proceed in a similar way. For example, we have

$$\begin{aligned}\widehat{\curlyvee}_\delta(n,t) &:= \int_0^t \sum_{\varepsilon_0\in\{-1,1\}} \frac{\varepsilon_0}{2i}\,\frac{e^{i\varepsilon_0\langle n\rangle(t-t')}}{\langle n\rangle}\widehat{\vee}_\delta(n,t')\,dt' \\ &= 2\int_{\mathbb{R}^2}\sum_{n_1,n_2\in\mathbb{Z}^3} Q^\delta_{n,n_1,n_2}(\tau_1,\tau_2)\,d\beta_{n_1}(\tau_1)\,d\beta_{n_2}(\tau_2),\end{aligned}$$

where

$$
\begin{aligned}
&Q^\delta_{n,n_1,n_2}(\tau_1,\tau_2)\\
&\quad = \int_{0\le t_1\le t_2\le t}\bigg(\int_{\tau_2}^{t}\sum_{\varepsilon_0,\varepsilon_1,\varepsilon_2\in\{-1,1\}}\frac{\varepsilon_0\varepsilon_1\varepsilon_2}{(2i)^3}\frac{e^{i\varepsilon_0\langle n\rangle(t-t')+i\sum_{j=1}^2\varepsilon_j\langle n_j\rangle(t'-t_j)}}{\langle n\rangle\langle n_1\rangle\langle n_2\rangle}\,dt'\bigg)\\
&\qquad\qquad\times\widehat{\rho}_\delta(n_1,t_1-\tau_1)\widehat{\rho}_\delta(n_2,t_2-\tau_2)\,dt_1\,dt_2.
\end{aligned}
$$

Note that the double Wiener–Itô integral accounts for the Wick renormalization on $\widehat{\mathbf{V}}_\delta$. Hence, we have

$$
\mathbb{E}[|\widehat{\Upsilon}_\delta(n,t)|^2]\sim\sum_{n_1,n_2\in\mathbb{Z}^3}\int_{\mathbb{R}^2}|Q^\delta_{n,n_1,n_2}(\tau_1,\tau_2)|^2\,d\tau_1\,d\tau_2.
$$

By Young's inequality in the two convolutions in time, we obtain

$$
\begin{aligned}
\mathbb{E}[|\widehat{\Upsilon}_\delta(n,t)|^2]\lesssim\sum_{n_1,n_2\in\mathbb{Z}^3}\int_{0\le t_1\le t_2\le t}\bigg|\int_{t_2}^{t}\sum_{\varepsilon_0,\varepsilon_1,\varepsilon_2\in\{-1,1\}}\varepsilon_0\varepsilon_1\varepsilon_2\\
\times\frac{e^{i\varepsilon_0\langle n\rangle(t-t')+i\sum_{j=1}^2\varepsilon_j\langle n_j\rangle(t'-t_j)}}{\langle n\rangle\langle n_1\rangle\langle n_2\rangle}\,dt'\bigg|^2\,dt_1\,dt_2
\end{aligned}
$$

uniformly in $0<\delta<1$ for the space-time mollifier ρ_δ. After integrating in t', the above expression essentially reduces to (3.16) in the proof of Proposition 1.6 and hence the rest follows as before. A more refined argument yields convergence in $L^p(\Omega)$ as the regularization is removed.

Recall that one of the key ingredients in studying the resonant product $\Upsilon\circledcirc\bullet$ and the paracontrolled operator $\mathfrak{I}_{\oslash,\ominus}$ is the symmetrization argument $n_2\leftrightarrow -n_2$, appearing in (4.14) and (5.17). It is at this point that we need to make use of the symmetry assumption (5.20). As in Remark 5.4, it suffices to consider

$$
\begin{aligned}
\sigma^\delta_{n_2}(t_1,t_2) &:= \mathbb{E}[\widehat{\uparrow}_\delta(n_2,t_1)\,\widehat{\uparrow}_\delta(-n_2,t_2)]\\
&= \mathbb{E}\big[\big(\rho_\delta(n_2,\cdot)*\widehat{\uparrow}(n_2,\cdot)\big)(t_1)\big(\rho_\delta(-n_2,\cdot)*\widehat{\uparrow}_\delta(-n_2,\cdot)\big)(t_2)\big]. \qquad (5.22)
\end{aligned}
$$

Using (5.21), we have

$$
\begin{aligned}
\sigma^\delta_{n_2}(t_1,t_2) &= -\int_0^{t_1}\int_0^{t_2}\sum_{\varepsilon_1,\varepsilon_2\in\{-1,1\}}\frac{\varepsilon_1\varepsilon_2}{4}\frac{e^{i\varepsilon_1\langle n_2\rangle(t_1-t_1')+i\varepsilon_2\langle n_2\rangle(t_2-t_2')}}{\langle n_2\rangle^2}\\
&\qquad\times\bigg(\int_{\mathbb{R}}\widehat{\rho}_\delta(n_2,t_1'-\tau)\widehat{\rho}_\delta(-n_2,t_2'-\tau)\,d\tau\bigg)\,dt_1'\,dt_2'\\
&= \sigma^\delta_{-n_2}(t_1,t_2),
\end{aligned}
$$

where we have used the symmetry assumption (5.20) in the last step. This shows that the effect of space-time mollifications in (5.22) is symmetric in $n_2\leftrightarrow -n_2$. This observation allows us to carry out the symmetrization argument in (4.14) and (5.17), provided that the space-time mollification kernel $\rho(x,t)$ is even in x.

6. Proof of Theorem 1.12

We now present the proof of Theorem 1.12. In the following, we assume that $0 < s_1 < s_2 < 1$. Recall that $(8, 8/3)$ and $(4, 4)$ are $\frac{1}{4}$-admissible and $\frac{1}{2}$-admissible, respectively. Given $0 < T \le 1$, we define $X_T^{s_1}$ (and $Y_T^{s_2}$) as the intersection of the energy spaces of regularity s_1 (and s_2, respectively) and the Strichartz space:

$$\begin{aligned} X_T^{s_1} &= C([0,T]; H^{s_1}(\mathbb{T}^3)) \cap C^1([0,T]; H^{s_1-1}(\mathbb{T}^3)) \cap L^8([0,T]; W^{s_1-1/4,8/3}(\mathbb{T}^3)), \\ Y_T^{s_2} &= C([0,T]; H^{s_2}(\mathbb{T}^3)) \cap C^1([0,T]; H^{s_2-1}(\mathbb{T}^3)) \cap L^4([0,T]; W^{s_2-1/2,4}(\mathbb{T}^3)), \end{aligned} \tag{6.1}$$

and set

$$Z_T^{s_1,s_2} = X_T^{s_1} \times Y_T^{s_2}.$$

Let $\Phi = (\Phi_1, \Phi_2)$ be as in (1.40) with the enhanced data set Ξ in (1.38) belonging to $\mathcal{X}_T^{s_1,s_2,\varepsilon}$ for some small $\varepsilon = \varepsilon(s_1, s_2) > 0$. By the Strichartz estimates (Lemma 2.4), the paraproduct estimate (Lemma 2.1), and the regularity assumptions on $\uparrow$ and $\curlyvee$, we have

$$\begin{aligned} \|\Phi_1(X,Y)\|_{X_T^{s_1}} &\lesssim \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + \|(X+Y-\curlyvee) \olessthan \uparrow\|_{L^1_T H^{s_1-1}} \\ &\lesssim \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + T\|X+Y-\curlyvee\|_{L^\infty_T L^2_x} \|\uparrow\|_{L^\infty_T W_x^{-1/2-\varepsilon,\infty}} \\ &\lesssim \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + T\big(1 + \|(X,Y)\|_{Z_T^{s_1,s_2}}\big) \end{aligned} \tag{6.2}$$

provided that $s_1 - 1 < -1/2 - \varepsilon$, that is $s_1 < 1/2$. Similarly, by Lemmas 2.4 and 2.1 with the regularity assumption on the enhanced data set Ξ in (1.38) and Corollary 5.2, we have

$$\begin{aligned} &\bigg\| S(t)(Y_0,Y_1) - \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} \big[2(X+Y-\curlyvee) \odot \uparrow + 2Y \circeq \uparrow - 2\curlyvee^{\circeq} \\ &\qquad\qquad + 2Z - 4\mathfrak{I}^{(1)}_{\olessthan}(X+Y-\curlyvee) \circeq \uparrow\big](t')\,dt'\bigg\|_{Y_T^{s_2}} \\ &\lesssim \|(Y_0,Y_1)\|_{\mathcal{H}^{s_2}} + \|(X+Y-\curlyvee)\odot\uparrow\|_{L^1_T H_x^{s_2-1}} + \|Y \circeq \uparrow\|_{L^1_T H_x^{s_2-1}} + \|\curlyvee^{\circeq}\|_{L^1_T H_x^{s_2-1}} \\ &\quad + \|Z\|_{L^1_T H_z^{s_2-1}} + \|\mathfrak{I}^{(1)}_{\olessthan}(X+Y-\curlyvee)\circeq\uparrow\|_{L^1_T H_x^{s_2-1}} \\ &\lesssim \|(Y_0,Y_1)\|_{\mathcal{H}^{s_2}} + T\big(1 + \|X+Y-\curlyvee\|_{L^\infty_T H_x^{s_1}} + \|Y\|_{L^\infty_T H_x^{s_2}}\big) \\ &\lesssim \|(Y_0,Y_1)\|_{\mathcal{H}^{s_2}} + T\big(1 + \|(X,Y)\|_{Z_T^{s_1,s_2}}\big) \end{aligned} \tag{6.3}$$

provided that $s_2 - 1 < \min(s_1 - 1/2 - 2\varepsilon, -\varepsilon)$ and $s_2 + (-1/2 - \varepsilon) > 0$, that is,

$$1/2 < s_2 < \min(1, s_1 + 1/2).$$

Similarly, we have

$$\begin{aligned} \bigg\| \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} \mathfrak{I}_{\olessthan,\circeq}(X+Y-\curlyvee)(t')\,dt' \bigg\|_{Y_T^{s_2}} &\lesssim \|\mathfrak{I}_{\olessthan,\circeq}(X+Y-\curlyvee)\|_{L^1_T H_x^{s_2-1}} \\ &\lesssim T\|X+Y-\curlyvee\|_{L^\infty_T L^2_x \cap C^1_T H_x^{-1-\varepsilon}} \lesssim T\big(1 + \|(X,Y)\|_{Z_T^{s_1,s_2}}\big) \end{aligned} \tag{6.4}$$

provided that $s_2 < 1$. Lastly, by Lemma 2.4 and the fractional Leibniz rule (Lemma 2.2), we have

$$\begin{aligned}\bigg\| \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}(X+Y-\curlyvee)^2(t')\,dt'\bigg\|_{Y_T^{s_2}} &\lesssim \|\langle\nabla\rangle^{s_2-1/2}(X+Y-\curlyvee)^2\|_{L^{4/3}_{T,x}}\\ &\lesssim T^{1/4}(\|\langle\nabla\rangle^{s_2-1/2}X\|^2_{L^8_T L^{8/3}_x}+\|\langle\nabla\rangle^{s_2-1/2}Y\|^2_{L^4_{T,x}}+\|\langle\nabla\rangle^{s_2-1/2}\curlyvee\|^2_{L^\infty_{T,x}})\\ &\lesssim T^{1/4}\big(1+\|(X,Y)\|^2_{Z_T^{s_1,s_2}}\big) \end{aligned} \tag{6.5}$$

provided that $s_2 \le \min(1-\varepsilon, s_1+1/4)$.

From (6.2)–(6.5), we obtain

$$\|\Phi(X,Y)\|_{Z_T^{s_1,s_2}} \lesssim \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + \|(Y_0,Y_1)\|_{\mathcal{H}^{s_2}} + T^\theta\big(1+\|(X,Y)\|^2_{Z_T^{s_1,s_2}}\big) \tag{6.6}$$

for some $\theta > 0$. By repeating a similar computation, we also obtain the following estimate of the difference:

$$\begin{aligned}&\|\Phi(X,Y)-\Phi(\widetilde{X},\widetilde{Y})\|_{Z_T^{s_1,s_2}}\\ &\quad\lesssim T^\theta\big(1+\|(X,Y)\|_{Z_T^{s_1,s_2}}+\|(\widetilde{X},\widetilde{Y})\|_{Z_T^{s_1,s_2}}\big)\|(X,Y)-(\widetilde{X},\widetilde{Y})\|_{Z_T^{s_1,s_2}}. \end{aligned}\tag{6.7}$$

Therefore, by choosing $T > 0$ sufficiently small (depending on the $\mathcal{X}_1^{s_1,s_2,\varepsilon}$-norm of the enhanced data set Ξ), we conclude from (6.6) and (6.7) that Φ in (1.40) is a contraction on the ball $B_R \subset Z_T^{s_1,s_2}$ of radius $R \sim \|(X_0,X_1)\|_{\mathcal{H}^{s_1}} + \|(Y_0,Y_1)\|_{\mathcal{H}^{s_2}}$. A similar computation yields continuous dependence of the solution (X,Y) on the enhanced data set Ξ measured in the $\mathcal{X}_1^{s_1,s_2,\varepsilon}$-norm. This concludes the proof of Theorem 1.12.

7. On the weak universality of the renormalized SNLW

We conclude this paper by presenting the proof of Theorem 1.2. For the sake of concreteness, we take the Gaussian noise η_κ to be the mollified space-time white noise $\rho * \xi$ on $(\kappa^{-1}\mathbb{T})^3 \times \mathbb{R}_+$ given by

$$\eta_\kappa = \kappa^{3/2}\sum_{n\in\mathbb{Z}^3}\widehat{\rho}(\kappa n)\frac{d\beta_n}{dt}e_{\kappa n}, \tag{7.1}$$

where ρ is a (smooth) mollification kernel with support in $\mathbb{T}^3 \cong [-1/2,1/2)^3$, $\{\beta_n\}_{n\in\Lambda_0}$ is a family of mutually independent complex-valued Brownian motions, and $\beta_{-n} := \overline{\beta_n}$, $n \in \Lambda_0$, as in (1.8). It is not difficult to see that η_κ is indeed a random field on $(\kappa^{-1}\mathbb{T})^3 \times \mathbb{R}_+$ which is smooth in space and white in time with stationary correlations.

Our aim is to describe the long time and large space behavior of the solution w_κ to (1.3). In order to do so, we perform a change of variables $u_\kappa(x,t) = \kappa^{-2}w_\kappa(\kappa^{-1}x,\kappa^{-1}t)$ as in (1.4). Then equation (1.3) takes the form

$$\partial_t^2 u_\kappa + (1-\Delta)u_\kappa = \kappa^{-4}f(\kappa^2 u_\kappa) + \kappa^{-4}a_\kappa^{(0)} + \kappa^{-2}a_\kappa^{(1)}u_\kappa + (1-\kappa^{-2})u_\kappa + \xi_\kappa \tag{7.2}$$

on $\mathbb{T}^3 \times \mathbb{R}_+$. Here, $\xi_\kappa(x,t) = \kappa^{-2}\eta_\kappa(\kappa^{-1}x, \kappa^{-1}t)$ is chosen so that ξ_κ converges in law to the space-time white noise ξ on $\mathbb{T}^3 \times \mathbb{R}_+$ as $\kappa \to 0$. Indeed, from (7.1), we deduce that

$$\xi_\kappa = \sum_{n\in\mathbb{Z}^3} \widehat{\rho}(\kappa n)\frac{d\widetilde{\beta}_n}{dt}e_n, \tag{7.3}$$

where $\{\widetilde{\beta}_n\}_{n\in\Lambda_0}$ is a family of mutually independent complex Brownian motions with the same joint law as $\{\beta_n\}_{n\in\Lambda_0}$ and $\widetilde{\beta}_{-n} := \overline{\widetilde{\beta}_n}$, $n \in \Lambda_0$. By taking

$$\xi = \sum_{n\in\mathbb{Z}^3} \frac{d\widetilde{\beta}_n}{dt}e_n$$

as a realization of the space-time white noise ξ, we see that ξ_κ converges to ξ in $C^{-1/2-\varepsilon}(\mathbb{R}_+; W^{-3/2-\varepsilon,\infty}(\mathbb{T}^3))$ (endowed with the compact-open topology) almost surely for any $\varepsilon > 0$.

By the Taylor remainder theorem, we can write the right-hand side of (7.2) (excluding ξ_κ) as

$$\begin{aligned}
&\kappa^{-4}f(\kappa^2 u_\kappa) + \kappa^{-4}a_\kappa^{(0)} + \kappa^{-2}a_\kappa^{(1)}u_\kappa + (1-\kappa^{-2})u_\kappa \\
&\quad = \{\kappa^{-4}f(0) + \kappa^{-4}a_\kappa^{(0)}\} + \{\kappa^{-2}f'(0) + \kappa^{-2}a_\kappa^{(1)} + (1-\kappa^{-2})\}u_\kappa + \frac{f''(0)}{2}u_\kappa^2 + R_\kappa,
\end{aligned}$$

where R_κ is the remainder given by

$$R_\kappa = \kappa^2 u_\kappa^3 \int_0^1 \frac{f'''(\tau\kappa^2 u_\kappa)}{6}(1-\tau)^2\, d\tau. \tag{7.4}$$

Let $\uparrow_\kappa$ be the solution of the linear equation

$$(\partial_t^2 + 1 - \Delta)\uparrow_\kappa = \xi_\kappa. \tag{7.5}$$

Then, with $b_\kappa(t) = \mathbb{E}[(\uparrow_\kappa(t))^2]$, we define the Wick power $\curlyvee_\kappa$ by

$$\curlyvee_\kappa = (\uparrow_\kappa)^2 - b_\kappa. \tag{7.6}$$

We now choose the time-dependent parameters $a_\kappa^{(0)}$ and $a_\kappa^{(1)}$ by setting

$$a_\kappa^{(0)} = -f(0) - \kappa^4 c_f b_\kappa \quad \text{and} \quad a_\kappa^{(1)} = -f'(0) + (1-\kappa^2), \tag{7.7}$$

where $c_f = f''(0)/2$. Then, by writing

$$u_\kappa = \uparrow_\kappa - w_\kappa,$$

we see from (7.2), (7.4), and (7.7) that v_κ satisfies

$$\partial_t^2 w_\kappa + (1-\Delta)w_\kappa = c_f\curlyvee_\kappa + 2c_f\uparrow_\kappa w_\kappa + c_f w_\kappa^2 + R_\kappa, \tag{7.8}$$

where we have used (7.6) to replace $(\uparrow_\kappa)^2 - b_\kappa$ by $\vee_\kappa$. In the following, by scaling, we assume that

$$c_f = -1.$$

Letting $\mathsf{Y}_\kappa = (\partial_t^2 + 1 - \Delta)^{-1}(\vee_\kappa)$, we decompose w_κ as

$$w_\kappa = -\mathsf{Y}_\kappa + X_\kappa + Y_\kappa$$

as in Section 1. Then, by repeating the discussion in Section 1, we can rewrite the equation (7.8) for w_κ into the following system for X_κ and Y_κ:

$$\begin{aligned} X_\kappa(t) &= -2\int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}[(X_\kappa + Y_\kappa - \mathsf{Y}_\kappa) \olessthan \uparrow_\kappa](t')\,dt', \\ Y_\kappa(t) &= -\int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}\Big[(X_\kappa + Y_\kappa - \mathsf{Y}_\kappa)^2 + 2(X_\kappa + Y_\kappa - \mathsf{Y}_\kappa) \ogreaterthan \uparrow_\kappa \\ &\qquad + 2Y_\kappa \oequal \uparrow_\kappa - 2\underset{\circ}{\mathsf{Y}}_\kappa - R_\kappa \\ &\qquad - 4\mathfrak{I}^{(1)}_{\olessthan}(X_\kappa + Y_\kappa - \mathsf{Y}_\kappa) \oequal \uparrow - 4\mathfrak{I}^{\kappa}_{\olessthan,\oequal}(X_\kappa Y_\kappa - \mathsf{Y}_\kappa)\Big](t')\,dt', \end{aligned} \tag{7.9}$$

where $\underset{\circ}{\mathsf{Y}}_\kappa = \mathsf{Y}_\kappa \oequal \uparrow$ and $\mathfrak{I}^{\kappa}_{\olessthan,\oequal}$ is defined as in (1.31) with $\uparrow$ replaced by $\uparrow_\kappa$.

Let $1/4 < s_1 < 1/2 < s_2 \le s_1 + 1/4$. Note that the rescaled noise ξ_κ in (7.3) is basically the mollified space-time white noise. Hence, it is easy to see that the enhanced data set associated with the rescaled noise ξ_κ

$$\Xi_\kappa = (0, 0, 0, 0, \uparrow_\kappa, \mathsf{Y}_\kappa, \underset{\circ}{\mathsf{Y}}_\kappa, 0, \mathfrak{I}^{\kappa}_{\olessthan,\oequal}), \tag{7.10}$$

belongs to the class $\mathcal{X}_1^{s_1,s_2,\varepsilon}$ defined in (1.39) since $\uparrow_\kappa$, Y_κ, $\underset{\circ}{\mathsf{Y}}_\kappa$, and $\mathfrak{I}^{\kappa}_{\olessthan,\oequal}$ satisfy the statements analogous to Lemma 3.1 and Propositions 1.6, 1.8, and 1.11.

Note that the system (7.9) is analogous to the original system (1.40) with the enhanced data set Ξ replaced by Ξ_κ and an additional source term given by the remainder term R_κ. In the following, we proceed as in the proof of Theorem 1.12 and prove local well-posedness of the system (7.9) for $\kappa > 0$ on a time interval $[0, T]$, where $T = T(\omega)$ is an almost surely positive stopping time, independent of $\kappa > 0$. Under the assumption $\|f'''\|_{L^\infty} < \infty$, we have $R_\kappa = O(\kappa^2 u_\kappa^3)$, where

$$u_\kappa = \uparrow_\kappa - \mathsf{Y}_\kappa + X_\kappa + Y_\kappa.$$

In order to handle the cubic structure of R_κ, we need to modify the norm for the second component Y_κ. Let $s_2 = 1/2 + \sigma$ with some small $\sigma > 0$. Noting that $(\frac{4}{1+2\sigma}, \frac{4}{1-2\sigma})$ is s_2-admissible, we define

$$\widetilde{Y}_T^{s_2} = C([0,T]; H^{s_2}(\mathbb{T}^3)) \cap C^1([0,T]; H^{s_2-1}(\mathbb{T}^3)) \cap L^{\frac{4}{1+2\sigma}}([0,T]; L^{\frac{4}{1-2\sigma}}(\mathbb{T}^3))$$

and set

$$\widetilde{Z}_T^{s_1,s_2} = X_T^{s_1} \times \widetilde{Y}_T^{s_2},$$

where $X_T^{s_1}$ is as in (6.1). In the following, we use the fact that $(\frac{4}{3+8\sigma}, \frac{4}{3-4\sigma})$ is dual s_2-admissible.

Note that (6.2)–(6.4) hold true even after replacing $Z_T^{s_1,s_2}$ and $Y_T^{s_2}$ by $\widetilde{Z}_T^{s_1,s_2}$ and $\widetilde{Y}_T^{s_2}$, respectively. Instead of (6.5), from the Strichartz estimates (Lemma 2.4) and Sobolev's inequality on X, we have

$$\begin{aligned}\bigg\| \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle}(X_\kappa + Y_\kappa - \curlyvee_\kappa)^2(t')\,dt'\bigg\|_{\widetilde{Y}_T^{s_2}} &\lesssim \|(X_\kappa + Y_\kappa - \curlyvee_\kappa)^2\|_{L_T^{\frac{4}{3+8\sigma}} L_x^{\frac{4}{3-4\sigma}}} \\ &\lesssim T^\theta\big(\|X_\kappa\|^2_{L_T^8 L_x^{\frac{8}{3-4\sigma}}} + \|Y_\kappa\|^2_{L_T^{\frac{4}{1+2\sigma}} L_x^{\frac{4}{1-2\sigma}}} + \|\curlyvee_\kappa\|^2_{L^\infty_{T,x}}\big) \\ &\le C(\omega)T^\theta\big(1 + \|(X_\kappa, Y_\kappa)\|^2_{\widetilde{Z}_T^{s_1,s_2}}\big) \qquad (7.11)\end{aligned}$$

for some $\theta > 0$, provided that $s_1 - 1/4 \ge 3\sigma/2$, which allows us to apply Sobolev's inequality

$$\|X_\kappa\|_{L_T^8 L_x^{\frac{8}{3-4\sigma}}} \lesssim \|X_\kappa\|_{L_T^8 W_x^{s_1-1/4,8/3}}.$$

Given $s_1 > 1/4$, this condition can be satisfied by choosing $\sigma > 0$ sufficiently small.

Next, we estimate the contribution from the remainder term R_κ. From (7.3) and (7.5), we see that $\widehat{\uparrow}_\kappa(n,t)$ is essentially supported on the spatial frequencies $\{|n| \lesssim \kappa^{-1}\}$. Hence, we have $\kappa^{1/2+\varepsilon}\uparrow_\kappa \in L^\infty([0,T]; L^\infty(\mathbb{T}^3))$ almost surely for any $\varepsilon > 0$. By a similar reasoning, the paracontrolled structure of the X_κ-equation in (7.9) allows us to conclude that X_κ essentially has the spatial frequency support in $\{|n| \lesssim \kappa^{-1}\}$. Therefore, from Lemma 2.4, (7.4) with $\|f'''\|_{L^\infty} < \infty$, and Sobolev's inequality, we have

$$\begin{aligned}\bigg\| \int_0^t \frac{\sin((t-t')\langle\nabla\rangle)}{\langle\nabla\rangle} R_\kappa(t')\,dt'\bigg\|_{\widetilde{Y}_T^{s_2}} &\lesssim \kappa^2\|(\uparrow_\kappa - \curlyvee_\kappa + X_\kappa + Y_\kappa)^3\|_{L_T^{\frac{4}{3+8\sigma}} L_x^{\frac{4}{3-4\sigma}}} \\ &\lesssim \kappa^{1/2-3\varepsilon}T(\|\kappa^{1/2+\varepsilon}\uparrow_\kappa\|^3_{L^\infty_{T,x}} + \|\curlyvee_\kappa\|^3_{L^\infty_{T,x}}) \\ &\quad + \kappa^{3\delta}T^\theta\big(\|\kappa^{2/3-\delta}X_\kappa\|^3_{L_T^8 L_x^{\frac{12}{3-4\sigma}}} + \|Y_\kappa\|^3_{L_T^{\frac{12}{3+8\sigma}} L_x^{\frac{12}{3-4\sigma}}}\big) \\ &\le C(\omega)\,\kappa^\delta T^\theta\big(1 + \|(X_\kappa, Y_\kappa)\|^3_{\widetilde{Z}_T^{s_1,s_2}}\big) \qquad (7.12)\end{aligned}$$

for some $\delta, \theta > 0$. Here we have used the frequency support of X_κ and Sobolev's inequality to bound

$$\|\kappa^{2/3-\delta}X_\kappa\|_{L_T^8 L_x^{\frac{12}{3-4\sigma}}} \lesssim \|X_\kappa\|_{L_T^8 W_x^{-2/3+\delta, \frac{12}{3-4\sigma}}} \lesssim \|X_\kappa\|_{L_T^8 W_x^{s_1-1/4,8/3}}$$

which holds when

$$3\left(\frac{3}{8} - \frac{3-4\sigma}{12}\right) = \frac{3}{8} + \sigma \le \left(s_1 - \frac{1}{4}\right) + \left(\frac{2}{3} - \delta\right) = s_1 + \frac{5}{12} - \delta.$$

This last condition is guaranteed by choosing $\sigma, \delta > 0$ sufficiently small. Note that we have used the bound

$$\|Y_\kappa\|_{L_T^{\frac{4}{1+8\sigma/3}} L_x^{\frac{4}{1-4\sigma/3}}} \lesssim T^{\sigma/6}\|Y_\kappa\|_{L_T^{\frac{4}{1+2\sigma}} L_x^{\frac{4}{1-2\sigma}}}.$$

Lastly, we point out that it was important to use s_2-admissible and dual s_2-admissible pairs such that there is no derivative on $\mathbb{R}_\kappa$ after applying the Strichartz estimate in (7.12).

Otherwise, a (fractional) derivative would fall on $f'''(\tau\kappa^2 u_\kappa)$ in (7.4) and we would need to use the fractional chain rule, which would make the computation far more complicated.

Putting (6.2), (6.3), (6.4), (7.11), and (7.12) together, we conclude that the system (7.9) is locally well-posed on $[0, T]$, where $T = T(\omega)$ is an almost surely positive stopping time, independent of $\kappa > 0$.

As for the sequence $\{\Xi_N\}_{N\in\mathbb{N}}$ above, one can show that, at least along subsequences, the family $\{\Xi_\kappa\}_{\kappa\in(0,1)}$ in (7.10) converges (in the natural $\mathcal{X}_1^{s_1,s_2,\varepsilon}$-topology) almost surely towards the random vector Ξ given by (1.42) with $(u_0, u_1) = (0, 0)$. Let (X, Y) be the solution to the original system (1.40) with this random data Ξ and define u by (1.43). Then, by using the above estimates, we can estimate the difference $(X - X_\kappa, Y - Y_\kappa)$. As a consequence, we conclude that, along any countable sequence, u_κ converges to the same limit u in $C([0, T]; H^{-1/2-\varepsilon}(\mathbb{T}^3))$ almost surely (and hence in probability), where $T = T(\omega)$ is a random local existence time whose size depends only on the random data Ξ, in particular, is independent of $\kappa \to 0$. Since the limit u does not depend on a particular countable sequence of $\kappa \to 0$, we can deduce that the whole family $\{u_\kappa\}_{\kappa\in(0,1)}$ converges in probability towards u. This completes the proof of Theorem 1.2.

Acknowledgments. The authors are grateful to the anonymous referee for the helpful comments that have improved the presentation of this paper.

Funding. M.G. and H.K. were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Hausdorff Center for Mathematics under Germany's Excellence Strategy–EXC–2047/1-390685813 and through CRC 1060–project number 211504053. T.O. was supported by the European Research Council (grant no. 637995 "ProbDynDispEq" and grant no. 864138 "SingStochDispDyn"). M.G. would like to thank the Isaac Newton Institute for Mathematical Sciences for support and hospitality during the programme "Scaling Limits, Rough Paths, Quantum Field Theory" when work on this paper was undertaken. This work was supported by EPSRC Grant Number EP/R014604/1.

References

[1] Albeverio, S., Haba, Z., Russo, F.: Trivial solutions for a non-linear two-space-dimensional wave equation perturbed by space-time white noise. Stochastics Stochastics Rep. **56**, 127–160 (1996) Zbl 0887.60069 MR 1396758

[2] Allez, R., Chouk, K.: The continuous Anderson hamiltonian in dimension two. arXiv:1511.02718 (2015)

[3] Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren Math. Wiss. 343, Springer, Heidelberg (2011) Zbl 1227.35004 MR 2768550

[4] Bailleul, I., Bernicot, F.: High order paracontrolled calculus. Forum Math. Sigma **7**, art. e44, 94 pp. (2019) Zbl 1473.60093 MR 4061967

[5] Bailleul, I., Debussche, A., Hofmanová, M.: Quasilinear generalized parabolic Anderson model equation. Stoch. Partial Differ. Equ. Anal. Comput. **7**, 40–63 (2019) Zbl 1448.35320 MR 3916262

[6] Bass, R. F.: Stochastic Processes. Cambridge Ser. Statist. Probab. Math. 33, Cambridge Univ. Press, Cambridge (2011) Zbl 1247.60001 MR 2856623

[7] 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) Zbl 1339.35281 MR 3350022

[8] Bényi, Á., Oh, T., Pocovnicu, O.: Higher order expansions for the probabilistic local Cauchy theory of the cubic nonlinear Schrödinger equation on $\mathbb{R}^3$. Trans. Amer. Math. Soc. Ser. B **6**, 114–160 (2019) Zbl 1410.35195 MR 3919013

[9] Bogachev, V. I.: Gaussian Measures. Math. Surveys Monogr. 62, Amer. Math. Soc., Providence, RI (1998) Zbl 0913.60035 MR 1642391

[10] Bony, J.-M.: Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. Ann. Sci. École Norm. Sup. (4) **14**, 209–246 (1981) Zbl 0495.35024 MR 631751

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

[12] 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

[13] Bruned, Y., Chandra, A., Chevyrev, I., Hairer, M.: Renormalising SPDEs in regularity structures. J. Eur. Math. Soc. **23**, 869–947 (2021) Zbl 1465.60057 MR 4210726

[14] Bruned, Y., Hairer, M., Zambotti, L.: Algebraic renormalisation of regularity structures. Invent. Math. **215**, 1039–1156 (2019) Zbl 1481.16038 MR 3935036

[15] 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

[16] 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

[17] Cannizzaro, G., Chouk, K.: Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. Ann. Probab. **46**, 1710–1763 (2018) Zbl 1407.60109 MR 3785598

[18] Cannizzaro, G., Matetski, K.: Space-time discrete KPZ equation. Comm. Math. Phys. **358**, 521–588 (2018) Zbl 1427.60118 MR 3774431

[19] Catellier, R., Chouk, K.: Paracontrolled distributions and the 3-dimensional stochastic quantization equation. Ann. Probab. **46**, 2621–2679 (2018) Zbl 1433.60048 MR 3846835

[20] Chandra, A., Hairer, M.: An analytic BPHZ theorem for regularity structures. arXiv:1612.08138 (2016)

[21] Chandra, A., Hairer, M., Shen, H.: The dynamical sine-Gordon model in the full subcritical regime. arXiv:1808.02594 (2018)

[22] Coifman, R. R., Meyer, Y.: Au delà des opérateurs pseudo-différentiels. Astérisque **57**, 185 pp. (1978) Zbl 0483.35082 MR 518170

[23] Da Prato, G., Debussche, A.: Two-dimensional Navier-Stokes equations driven by a space-time white noise. J. Funct. Anal. **196**, 180–210 (2002) Zbl 1013.60051 MR 1941997

[24] Da Prato, G., Debussche, A.: Strong solutions to the stochastic quantization equations. Ann. Probab. **31**, 1900–1916 (2003) Zbl 1071.81070 MR 2016604

[25] Debussche, A., Martin, J.: Solution to the stochastic Schrödinger equation on the full space. Nonlinearity **32**, 1147–1174 (2019) Zbl 1407.35236 MR 3923163

[26] Debussche, A., Weber, H.: The Schrödinger equation with spatial white noise potential. Electron. J. Probab. **23**, art. 28, 16 pp. (2018) Zbl 1387.60097 MR 3785398

[27] Deng, Y., Nahmod, A., Yue, H.: Invariant Gibbs measures and global strong solutions for nonlinear Schrödinger equations in dimension two. arXiv:1910.08492 (2019)

[28] 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

[29] Fan, D., Sato, S.: Transference on certain multilinear multiplier operators. J. Austral. Math. Soc. **70**, 37–55 (2001) Zbl 0984.42006 MR 1808390

[30] Friz, P. K., Hairer, M.: A Course on Rough Paths. Universitext, Springer, Cham (2014) Zbl 1327.60013 MR 3289027

[31] Funaki, T., Hoshino, M.: A coupled KPZ equation, its two types of approximations and existence of global solutions. J. Funct. Anal. **273**, 1165–1204 (2017) Zbl 1370.60104 MR 3653951

[32] Furlan, M., Gubinelli, M.: Paracontrolled quasilinear SPDEs. Ann. Probab. **47**, 1096–1135 (2019) Zbl 1447.60099 MR 3916943

[33] Furlan, M., Gubinelli, M.: Weak universality for a class of 3d stochastic reaction-diffusion models. Probab. Theory Related Fields **173**, 1099–1164 (2019) Zbl 1411.60095 MR 3936152

[34] Gerencsér, M., Hairer, M.: A solution theory for quasilinear singular SPDEs. Comm. Pure Appl. Math. **72**, 1983–2005 (2019) Zbl 1458.60077 MR 3987723

[35] Ginibre, J., Velo, G.: Generalized Strichartz inequalities for the wave equation. J. Funct. Anal. **133**, 50–68 (1995) Zbl 0849.35064 MR 1351643

[36] Gubinelli, M.: Rough solutions for the periodic Korteweg–de Vries equation. Comm. Pure Appl. Anal. **11**, 709–733 (2012) Zbl 1278.35213 MR 2861805

[37] Gubinelli, M., Hofmanová, M.: Global solutions to elliptic and parabolic Φ^4 models in Euclidean space. Comm. Math. Phys. **368**, 1201–1266 (2019) Zbl 1420.35481 MR 3951704

[38] Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. Forum Math. Pi **3**, art. e6, 75 pp. (2015) Zbl 1333.60149 MR 3406823

[39] Gubinelli, M., Koch, H., Oh, T.: Renormalization of the two-dimensional stochastic nonlinear wave equations. Trans. Amer. Math. Soc. **370**, 7335–7359 (2018) Zbl 1400.35240 MR 3841850

[40] Gubinelli, M., Koch, H., Oh, T., Tolomeo, L.: Global dynamics for the two-dimensional stochastic nonlinear wave equations. Int. Math. Res. Notices **2022**, 16954–16999 Zbl 07614307 MR 4504911

[41] Gubinelli, M., Perkowski, N.: KPZ reloaded. Comm. Math. Phys. **349**, 165–269 (2017) Zbl 1388.60110 MR 3592748

[42] Gubinelli, M., Ugurcan, B., Zachhuber, I.: Semilinear evolution equations for the Anderson Hamiltonian in two and three dimensions. Stoch. Partial Differ. Equ. Anal. Comput. **8**, 82–149 (2020) Zbl 1436.35331 MR 4058957

[43] Hairer, M.: Solving the KPZ equation. Ann. of Math. (2) **178**, 559–664 (2013) Zbl 1281.60060 MR 3071506

[44] Hairer, M.: Singular stochastic PDEs. In: Proc. Int. Congress of Mathematicians—Seoul 2014. Vol. 1, Kyung Moon Sa, Seoul, 685–709 (2014) Zbl 1373.60110 MR 3728488

[45] Hairer, M.: A theory of regularity structures. Invent. Math. **198**, 269–504 (2014) Zbl 1332.60093 MR 3274562

[46] Hairer, M.: Renormalisation of parabolic stochastic PDEs. Jpn. J. Math. **13**, 187–233 (2018) Zbl 1430.60055 MR 3855740

[47] Hairer, M.: The motion of a random string. arXiv:1605.02192 (2016)

[48] Hairer, M., Labbé, C.: Multiplicative stochastic heat equations on the whole space. J. Eur. Math. Soc. **20**, 1005–1054 (2018) Zbl 1447.60102 MR 3779690

[49] Hairer, M., Matetski, K.: Discretisations of rough stochastic PDEs. Ann. Probab. **46**, 1651–1709 (2018) Zbl 1406.60094 MR 3785597

[50] Hairer, M., Quastel, J.: A class of growth models rescaling to KPZ. Forum Math. Pi **6**, art. e3, 112 pp. (2018) Zbl 1429.60057 MR 3877863

[51] Hairer, M., Shen, H.: The dynamical sine-Gordon model. Comm. Math. Phys. **341**, 933–989 (2016) Zbl 1336.60120 MR 3452276

[52] Hairer, M., Shen, H.: A central limit theorem for the KPZ equation. Ann. Probab. **45**, 4167–4221 (2017) Zbl 1388.60111 MR 3737909

[53] Hairer, M., Xu, W.: Large-scale behavior of three-dimensional continuous phase coexistence models. Comm. Pure Appl. Math. **71**, 688–746 (2018) Zbl 1407.60085 MR 3772400

[54] Hairer, M., Xu, W.: Large scale limit of interface fluctuation models. Ann. Probab. **47**, 3478–3550 (2019) Zbl 1453.60120 MR 4038037

[55] Hoshino, M.: Global well-posedness of complex Ginzburg–Landau equation with a space-time white noise. Ann. Inst. H. Poincaré Probab. Statist. **54**, 1969–2001 (2018) Zbl 1418.60079 MR 3865664

[56] Hoshino, M.: Paracontrolled calculus and Funaki–Quastel approximation for the KPZ equation. Stochastic Process. Appl. **128**, 1238–1293 (2018) Zbl 1384.35149 MR 3769661

[57] Hoshino, M., Inahama, Y., Naganuma, N.: Stochastic complex Ginzburg–Landau equation with space-time white noise. Electron. J. Probab. **22**, art. 104, 68 pp. (2017) Zbl 1386.60218 MR 3742401

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

[59] Killip, R., Stovall, B., Visan, M.: Blowup behaviour for the nonlinear Klein–Gordon equation. Math. Ann. **358**, 289–350 (2014) Zbl 1290.35227 MR 3157999

[60] Kupiainen, A.: Renormalization group and stochastic PDEs. Ann. Henri Poincaré **17**, 497–535 (2016) Zbl 1347.81063 MR 3459120

[61] Kupiainen, A., Marcozzi, M.: Renormalization of generalized KPZ equation. J. Statist. Phys. **166**, 876–902 (2017) Zbl 1369.82011 MR 3607594

[62] Lindblad, H., Sogge, C. D.: On existence and scattering with minimal regularity for semilinear wave equations. J. Funct. Anal. **130**, 357–426 (1995) Zbl 0846.35085 MR 1335386

[63] 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

[64] Martin, J., Perkowski, N.: Paracontrolled distributions on Bravais lattices and weak universality of the 2d parabolic Anderson model. Ann. Inst. H. Poincaré Probab. Statist. **55**, 2058–2110 (2019) Zbl 1446.60047 MR 4029148

[65] Matetski, K.: Martingale-driven approximations of singular stochastic PDEs. arXiv:1808.09429 (2018)

[66] McKean, H. P.: Statistical mechanics of nonlinear wave equations. IV. Cubic Schrödinger. Comm. Math. Phys. **168**, 479–491 (1995) Zbl 0821.60069 MR 1328250

[67] Mourrat, J.-C., Weber, H.: The dynamic Φ^4_3 model comes down from infinity. Comm. Math. Phys. **356**, 673–753 (2017) Zbl 1384.81068 MR 3719541

[68] Mourrat, J.-C., Weber, H.: Global well-posedness of the dynamic Φ^4 model in the plane. Ann. Probab. **45**, 2398–2476 (2017) Zbl 1381.60098 MR 3693966

[69] Mourrat, J.-C., Weber, H., Xu, W.: Construction of Φ^4_3 diagrams for pedestrians. In: From Particle Systems to Partial Differential Equations, Springer Proc. Math. Statist. 209, Springer, Cham, 1–46 (2017) Zbl 1390.81266 MR 3746744

[70] Muscalu, C., Schlag, W.: Classical and Multilinear Harmonic Analysis. Vol. II. Cambridge Stud. Adv. Math. 138, Cambridge Univ. Press, Cambridge (2013) Zbl 1281.42002 MR 3052499

[71] Nelson, E.: A quartic interaction in two dimensions. In: Mathematical Theory of Elementary Particles (Dedham, MA, 1965), M.I.T. Press, Cambridge, MA, 69–73 (1966) MR 0210416

[72] Oberguggenberger, M., Russo, F.: Nonlinear stochastic wave equations. Integral Transforms Special Funct. **6**, 71–83 (1998) Zbl 0912.60072 MR 1640497

[73] Oberguggenberger, M., Russo, F.: Singular limiting behavior in nonlinear stochastic wave equations. In: Stochastic Analysis and Mathematical Physics, Progr. Probab. 50, Birkhäuser Boston, Boston, MA, 87–99 (2001) Zbl 0987.60075 MR 1886565

[74] Oh, T.: Periodic stochastic Korteweg–de Vries equation with additive space-time white noise. Anal. PDE **2**, 281–304 (2009) Zbl 1190.35202 MR 2603800

[75] Oh, T., Okamoto, M., Robert, T.: A remark on triviality for the two-dimensional stochastic nonlinear wave equation. Stochastic Process. Appl. **130**, 5838–5864 (2020) Zbl 1448.35587 MR 4127348

[76] Oh, T., Okamoto, M., Tzvetkov, N.: Uniqueness and non-uniqueness of the Gaussian free field evolution under the two-dimensional Wick ordered cubic wave equation. arXiv:2206.00728 (2022)

[77] 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

[78] Oh, T., Pocovnicu, O.: A remark on almost sure global well-posedness of the energy-critical defocusing nonlinear wave equations in the periodic setting. Tohoku Math. J. (2) **69**, 455–481 (2017) Zbl 1394.35285 MR 3695994

[79] Oh, T., Robert, T., Wang, Y.: On the parabolic and hyperbolic Liouville equations. Comm. Math. Phys. **387**, 1281–1351 (2021) Zbl 1481.60123 MR 4324379

[80] Oh, T., Thomann, L.: Invariant Gibbs measures for the 2-d defocusing nonlinear wave equations. Ann. Fac. Sci. Toulouse Math. (6) **29**, 1–26 (2020) Zbl 1443.35101 MR 4133695

[81] Oh, T., Tzvetkov, N., Wang, Y.: Solving the 4NLS with white noise initial data. Forum Math. Sigma **8**, art. e48, 63 pp. (2020) Zbl 1452.35193 MR 4176752

[82] Oh, T., Robert, T., Tzvetkov, N.: Stochastic nonlinear wave dynamics on compact surfaces. Ann. H. Lebesgue, to appear

[83] Otto, F., Sauer, J., Smith, S., Weber, H.: Parabolic equations with rough coefficients and singular forcing. arXiv:1803.07884 (2018)

[84] Otto, F., Weber, H.: Quasilinear SPDEs via rough paths. Arch. Ration. Mech. Anal. **232**, 873–950 (2019) Zbl 1426.60090 MR 3925533

[85] Perkowski, N., Rosati, T. C.: The KPZ equation on the real line. Electron. J. Probab. **24**, art. 117, 56 pp. (2019) Zbl 1427.60132 MR 4029420

[86] 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

[87] Russo, F.: Colombeau generalized functions and stochastic analysis. In: Stochastic Analysis and Applications in Physics (Funchal, 1993), NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci. 449, Kluwer, Dordrecht, 329–349 (1994) MR 1337971

[88] Shen, H.: Stochastic quantization of an Abelian gauge theory. Comm. Math. Phys. **384**, 1445–1512 (2021) Zbl 07356544 MR 4268826

[89] Shen, H., Weber, H.: Glauber dynamics of 2D Kac–Blume–Capel model and their stochastic PDE limits. J. Funct. Anal. **275**, 1321–1367 (2018) Zbl 1395.82137 MR 3820327

[90] Shen, H., Xu, W.: Weak universality of dynamical Φ^4_3: non-Gaussian noise. Stoch. Partial Differ. Equ. Anal. Comput. **6**, 211–254 (2018) Zbl 1400.37101 MR 3818405

[91] Shigekawa, I.: Stochastic Analysis. Transl. Math. Monogr. 224, Amer. Math. Soc., Providence, RI (2004) Zbl 1064.60003 MR 2060917

[92] Simon, B.: The $P(\phi)_2$ Euclidean (Quantum) Field Theory. Princeton Ser. Phys., Princeton Univ. Press, Princeton, NJ (1974) Zbl 1175.81146 MR 0489552

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

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

[95] Tolomeo, L.: Global well posedness of the two-dimensional stochastic nonlinear wave equation on an unbounded domain. Ann. Probab. **49**, 1402–1426 (2021) Zbl 1467.35221 MR 4255148

[96] Zhu, R., Zhu, X.: Three-dimensional Navier–Stokes equations driven by space-time white noise. J. Differential Equations **259**, 4443–4508 (2015) Zbl 1336.60127 MR 3373412