2D-defocusing nonlinear Schrödinger equation with random data on irrational tori

Chenjie Fan, Yumeng Ou, Gigliola Staffilani & Hong Wang

Stochastics and Partial Differential Equations: Analysis and Computations

ISSN 2194-0401

Stoch PDE: Anal Comp
DOI 10.1007/s40072-020-00174-7

Springer

Stoch PDE: Anal Comp
https://doi.org/10.1007/s40072-020-00174-7

2D-defocusing nonlinear Schrödinger equation with random data on irrational tori

Chenjie Fan[1] · Yumeng Ou[2] · Gigliola Staffilani[3] · Hong Wang[4]

Received: 27 October 2019 / Revised: 11 April 2020

Abstract
We revisit the work of Bourgain on the invariance of the Gibbs measure for the cubic, defocusing nonlinear Schrödinger equation in 2D on a square torus, and we prove the equivalent result on any tori.

Keywords Random Data · NLS · LWP

1 Introduction

The purpose of this work is to revisit the famous work of Bourgain on the invariance of the Gibbs measure for the 2D defocusing cubic nonlinear Schrödinger equation (NLS) on a square torus $\mathbb{T}^2$, [4], and extend his proof to any torus. Since later we often use the definition of *rational* or *irrational* torus, we readily give it here. Assume

C.F. is funded in part by an AMS-Simons Foundation travel grant. Y.O. is funded in part by NSF DMS-1854148. G.S. is funded in part by NSF DMS-1462401 and DMS-1764403, and the Simons Foundation. H.W. is funded by the S.S. Chern Foundation for Mathematics Research Fund and the National Science Foundation.

✉ Chenjie Fan
cjfanpku@gmail.com

Yumeng Ou
yumeng.ou@baruch.cuny.edu

Gigliola Staffilani
gigliola@math.mit.edu

Hong Wang
hongwang@ias.edu

[1] Department of Mathematics, University of Chicago, Chicago, IL, USA

[2] Department of Mathematics, Baruch College, City University of New York, New York, NY, USA

[3] Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

[4] Institute for Advanced Study, Princeton, NJ, USA

Published online: 07 May 2020

that a 2D torus $\mathbb{T}^2$ has periods T_1 and T_2. If T_1/T_2 is rational we call $\mathbb{T}^2$ a rational torus, otherwise irrational. As one can see from the proof for Strichartz estimates on rational tori [2], Bourgain uses a fundamental property of the linear solution of the Schrödinger equation defined on a rational torus, the fact that the solution is also periodic in time. In the proof this fact is used in reducing the Strichartz inequality to estimating the cardinality of a set of lattice points $(x, y) \in \mathbb{Z}^2$ that satisfy a quadratic equation $x^2+ay^2 = R^2$, where a, R are natural numbers. Then an well known theorem from number theory is invoked to give a sharp bound in terms of R, see [1], see also (2.43), page 117 in [2]. If one wants to repeat Bourgain's proof for generic tori one has to obtain the same sharp bound when counting the lattice points in a region such as $\{(x, y) \ / \ x^2 + \alpha y^2 \leq R^2 + O(1)\}$, where now γ, $R^2 > 0$. In general, this number of lattice points is larger than the sharp bound above. Intuitively though the same type of Strichartz estimates available for rational tori should be available for irrational one. In fact one expects even better ones since the irrationality of the torus should translate into fewer interactions among linear Schrödinger solutions. After almost two decades the full range of Strichartz estimates on any torus were proved by Bourgain and Demeter [5], who obtained them as a corollary of their proof of the l^2 *decoupling conjecture*, hence without invoking any number theory. Shortly later Deng, Germain and Guth [7] proved that indeed Strichartz estimates on a *generic* irrational torus are better, in the sense that they *live* for a longer interval of time.

Let us now go back to Bourgain's work on the invariance of the Gibbs measure in [4]. If one considers the nonlinear Schrödinger equation with solution u and Hamiltonian $H(u)$ in the frequency space instead of the physical space, then it can be recast as an infinite dimensional Hamiltonian system with variables $(q_n(t), p_n(t))$ such that $\hat{u}(t, n) = q_n(t) + ip_n(t)$ for frequencies $n \in \mathbb{Z}^2$. For this infinite dimensional system one can define a Gibbs measure that formally can be written as $d\mu = 1/Ze^{H(u)}du$, where Z is a normalizing factor to make it a probability measure, and its support lives in $H^s(\mathbb{T}^2)$, $s < 0$, see [9,10]. Bourgain had already proved [3] that for the 1D quintic NLS, where a similar measure can be defined with support in $H^s(\mathbb{T})$, $s < 1/2$ [10], the Schrödinger flow indeed leaves the measure invariant, meaning for any set A in the support of the measure, its evolution with respect to the Schrödinger flow at any later times has the same measure as A itself. Moreover using this invariance he proved that the flow can be defined globally almost surely. Clearly such a question could have been asked in 2D for the cubic defocusing NLS[1] as well. The issue that Bourgain faced was that while in the 1D case the flow was (deterministically) defined, at least locally, for any data in $H^s(\mathbb{T})$, $s > 0$, and hence on the whole support of the Gibbs measure, which as recalled is in H^s, $s < 1/2$, for the cubic 2D case also the flow was only known to be defined for data in $H^s(\mathbb{T}^2)$, $s > 0$, hence missing the support of the Gibbs measure, which is in H^s, $s < 0$. To overcome this, and other serious analytic obstacles along the way, Bourgain used probabilistic tools, such as Wick ordering and large deviation estimates, combined with more deterministic ones, such as Strichartz type estimates and counting lemmata similar to the ones recalled above. This brings us to the motivation of our paper. Indeed here we prove new counting lemmata, see

[1] It is known that while for the 1D quintic focusing NLS the Gibbs measure can be defined as long as the L^2 norm is smaller than a certain absolute constant, in 2D no Gibbs measure can be defined for the focusing case [10].

Sect. 4, that hold more generally for any torus, and we rework the local almost sure well-posedness in details[2] to show that the bounds obtained in the counting lemmata, although weaker than the ones in [4] still are enough to conclude the proof. Although our paper follows the scheme of Bourgain's proof, we decided to add full details for the convenience of the reader and because along the way we could point out with remarks where special care needed to be put in place in order to treat the general case and rationality cannot be invoked.

Finally, we recently learned that Deng, Nahmod and Yue [8] have extended Bourgain's result in [4] for any nonlinearity $2r+1$, where $r \in \mathbb{N}$. This is a remarkable feat since the high order of nonlinear interactions was previously considered an almost insurmountable obstacle in obtaining an almost sure local flow in the support of the Gibbs measure.

1.1 Statement of main result

In this paper, we study the 2D cubic Wick ordered NLS equation on irrational or rational tori. We will pose the NLS on a rectangular torus and rescale the Δ. Let $\gamma \in (1,2)$ be any real number (possibly irrational) that determines on $\mathbb{T}^2 = [0, 2\pi]^2$ the operator

$$\Delta_\gamma := \partial_x^2 + \frac{1}{\gamma}\partial_y^2.$$

The free solution to the linear Schrödinger initial value problem

$$\begin{cases} iu_t - \Delta_\gamma u = 0, \\ u_0 = \sum_{n\in\mathbb{Z}^2} a_n e^{in\cdot x} \end{cases} \tag{1.1}$$

is of the form

$$S(t)u_0 \equiv e^{it\Delta_\gamma}u_0 \equiv \sum_{n\in\mathbb{Z}^2} a_n e^{in\cdot x} e^{in^2 t},$$

where we let

$$n^2 := n\cdot n = \langle n, n\rangle_\gamma = n_1^2 + \gamma n_2^2.$$

Following the set up of Bourgain [4] we revisit the Wick ordered truncated NLS with random initial data

$$\begin{cases} i\partial_t u_N - \Delta_\gamma u_N = -P_{\le N}[(|u_N|^2 - M_N)u_N], \\ u_{0,N} = \sum_{|n|\le N} \frac{g_n(\omega)}{|n|} e^{in\cdot x}, \ x \in \mathbb{T}^2, \end{cases} \tag{1.2}$$

[2] Here we will not repeat the argument that upgrades the local well-posedness to the global since the rationality or not of the torus plays no role.

where $\{g_n(\omega) : n \in \mathbb{Z}^2\}$ are independent L^2-normalized complex Gaussians and

$$M_N := 2\int_{\mathbb{T}^2} |u_N(t,x)|^2\,dx = 2\int_{\mathbb{T}^2} |u_{0,N}|^2.$$

The operator $P_{\le N}$ here denotes the projection onto frequencies $|n| \le N$. By definition of M_N, one may rewrite (1.2) as

$$\begin{cases} i\partial_t u_N - \Delta_\gamma u_N = -P_{\le N}\mathcal{N}(u_N), \\ u_{0,N} = \sum_{|n|\le N} \frac{g_n(\omega)}{|n|} e^{in\cdot x},\ x \in \mathbb{T}^2 \end{cases} \tag{1.3}$$

by defining the Wick ordered nonlinearity

$$\begin{aligned} \mathcal{N}(f,g,h)(x) = &\sum_{n_i\in\mathbb{Z}^2,\, n_2\neq n_1,n_3} \widehat{f}(n_1)\overline{\widehat{g}(n_2)}\widehat{h}(n_3)e^{i\langle n_1-n_2+n_3,x\rangle} \\ &- \sum_{n\in\mathbb{Z}^2} \widehat{f}(n)\overline{\widehat{g}(n)}\widehat{h}(n)e^{in\cdot x} \\ =:& \mathcal{N}_1(f,g,h) + \mathcal{N}_2(f,g,h), \end{aligned} \tag{1.4}$$

and we write $\mathcal{N}(f) := \mathcal{N}(f,f,f)$, similarly for $\mathcal{N}_i$, $i = 1, 2$.

One may also study the formal limit equation of (1.3)

$$\begin{cases} i\partial_t u - \Delta_\gamma u = -\mathcal{N}(u), \\ u_0 = \sum_n \frac{g_n(\omega)}{|n|} e^{in\cdot x},\ x \in \mathbb{T}^2. \end{cases} \tag{1.5}$$

For every ω and every N fixed, equation (1.3) is finite dimensional and thus an ODE. It hence has a local solution, and in fact also a global solution due to the mass conservation law. Therefore, one is mainly interested in a local theory for (1.3) that is independent of N. More precisely, the main result of the paper is the following

Theorem 1.1 *Let u_N^ω be the solution to (1.3),*

$$u_N^\omega(t,x) = e^{it\Delta_\gamma}u_{0,N} - i\int_0^t e^{i(t-s)\Delta_\gamma} P_{\le N}\mathcal{N}(u_N)\,ds, \quad \forall 0 \le t < t_\omega. \tag{1.6}$$

There exists $s_0 > 0, \epsilon_0 > 0$, so that for almost every $\omega \in \Omega$, there exists a t_ω independent of N, such that

$$\|u_N^\omega(t,x) - e^{it\Delta_\gamma}u_{0,N}\|_{X^{s_0,\frac{1}{2}+\epsilon_0}} \lesssim 1. \tag{1.7}$$

Moreover, for any $0 < s' < s_0$, $w_N^\omega := u_N^\omega(t,x) - e^{it\Delta_\gamma}u_{0,N}$ converges strongly in $X^{s',\frac{1}{2}+\epsilon_0}[0,t_\omega]$ to some limit w. Furthermore, the limit $u^\omega := w + e^{i\Delta t}u_0$ (called the solution to the Wick ordered NLS (1.5)) satisfies the Duhamel formula

$$u^{\omega}(t,x) = e^{it\Delta_\gamma} u_0 - i\int_0^t e^{i(t-s)\Delta_\gamma}\mathcal{N}(u^{\omega})\,ds, \quad \forall 0 \le t < t_\omega. \tag{1.8}$$

We refer the reader to Sect. 3.1 for the definition of the $X^{s,b}$ space and its related properties. In the following, when the dependence on the parameter $\omega \in \Omega$ is clear from the context, we sometimes will drop the superscript in u^{ω} and write u for short.

Theorem 1.1 follows from the following quantitative version (independent of N) of the main result.

Theorem 1.2 *Let u_N^{ω} be as in Theorem 1.1, there exists $s_0 > 0, \epsilon_0, \alpha_0, t_0 > 0$ so that for every $t < t_0$, up to a set of probability measure $e^{-t^{-\alpha_0}}$, one has that $w_N := u_N - e^{it\Delta}u_{0,N}$ satisfies $\|w_N\|_{X^{s_0,\frac{1}{2}+\epsilon_0}[0,t]} \lesssim 1$. Furthermore, for all $0 < s' < s_0$, w_N converges to some limit w in $X^{s',\frac{1}{2}+\epsilon_0}$ and $u = w + e^{it\Delta_\gamma}u_0$ solves the wick ordered NLS in the sense that Duhamel formula (1.8) is satisfied.*

1.2 Duhamel formula, Picard iteration and main propositions

The proof follows from a Picard iteration scheme. One would like to write (1.3) into its Duhamel form, (1.6). It is convenient to introduce an extra time cut off $\phi_\delta(t) = \phi(t/\delta)$, where $\phi \in C_0^\infty(\mathbb{R})$ is equal to 1 on $[-1/2, 1/2]$ and 0 outside $[-1, 1]$, and consider instead the following slightly modified version of (1.6):

$$u_N'(t,x) = \phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} - i\phi_\delta(t)\int_0^t e^{i(t-s)\Delta_\gamma}(P_{\le N}\mathcal{N}(\phi(t)u_N'))\,ds. \tag{1.9}$$

Note that when $t < \delta/2$, u_N' is no different from u_N. In what follows, for convenience we will not distinguish u_N and u_N'. Let

$$\Gamma_{N,\delta}u = -i\phi_\delta(t)\int_0^t e^{i(t-s)\Delta_\gamma}P_{\le N}\mathcal{N}(\phi_\delta(t)u)\,ds, \tag{1.10}$$

and consider its formal limit as $N \to \infty$:

$$\Gamma_\delta u = i\phi_\delta(t)\int_0^t e^{i(t-s)\Delta_\gamma}\mathcal{N}(\phi_\delta(t)u)\,ds. \tag{1.11}$$

From the perturbative viewpoint, let

$$u_N = \phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} + w_N(x,t). \tag{1.12}$$

Then (1.9) is equivalent to

$$w_N(x,t) = \Gamma_{N,\delta}(\phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} + w_N(x,t)), \tag{1.13}$$

which reduces Theorem 1.2 to the following three propositions,

Proposition 1.3 *There exists a sufficiently small $\delta_0 > 0$ and $s_0 \gg \epsilon_0 > 0$, and some $\alpha_0 > 0$, such that for every $0 < \delta < \delta_0$, up to a set of measure $e^{-\delta^{-\alpha_0}}$ for some α_0 depending on ϵ_0, the map*

$$w \mapsto \Gamma_{N,\delta}(\phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} + w(x,t)) \tag{1.14}$$

is a contraction map on the space

$$\{w : \|w\|_{X^{s_0,b_0}} \leq 1\} \text{ for all } N. \tag{1.15}$$

where $b_0 = 1/2 + \epsilon_0$.

Proposition 1.4 *There exists a sufficiently small $\delta_0 > 0$ and $s_0 \gg \epsilon_0 > 0$, and $\alpha_0 > 0$ s.t. for every $0 < \delta < \delta_0$, up to a set of measure $e^{-\delta^{-\alpha_0}}$ for some α_0 depending on ϵ_0, the map*

$$w \mapsto \Gamma_{\delta}(\phi_\delta(t)e^{it\Delta_\gamma}u_{0} + w(x,t)) \tag{1.16}$$

is a contraction map on the space

$$\{w : \|w\|_{X^{s_0,b_0}} \leq 1\}, \tag{1.17}$$

where $b_0 = 1/2 + \epsilon_0$.

Proposition 1.5 *Let δ_0, s_0, ϵ_0 be as in Proposition 1.3 and 1.4. Let w_N be the unique function (fixed point) in $\{w : \|w\|_{X^{s_0,b_0}} \leq 1\}$ such that*

$$w_N = \Gamma_{N,\delta}(\phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} + w_N(x,t)), \tag{1.18}$$

and let w^* be the unique function (fixed point) in $\{w : \|w\|_{X^{s_0,b_0}} \leq 1\}$ such that

$$w^* = \Gamma_{\delta}(\phi_\delta(t)e^{it\Delta_\gamma}u_{0,N} + w^*(x,t)). \tag{1.19}$$

Then one has for all $s' < s_0$ and as $N \to \infty$ that

$$w_N \to w^* \text{ in } X^{s',\frac{1}{2}+\epsilon_0}. \tag{1.20}$$

Remark 1.6 Note that Proposition 1.3 is stated uniformly over all $N > 0$. In particular, it can be seen from the proof that, the exceptional set, does not depend on N. We are following [4] here, see also discussions in [8,13]. Thus to prove Proposition 1.3 is equivalent to prove Proposition 1.4. For those who are familiar with the Picard iteration scheme, Proposition 1.5 is a stability argument that is essentially equivalent to the local existence argument giving Proposition 1.3 and 1.4. However, to take into account the difference between $P_{\leq N}\mathcal{N}(u_N)$ and $\mathcal{N}(u_N)$, one will need to use extra derivative, which is the reason why the convergence in Proposition 1.5 only holds for $s' < s_0$. One may also see [8],

We will focus on the proof of Proposition 1.4, which is the same as the one for Proposition 1.3, then Proposition 1.5 follows by using the same argument as in Section 5 of [4]. One may also use the invariance of the Gibbs measure to upgrade the local well-posedness to a global one as in [4].

2 Proof of Proposition 1.4: initial reduction and structure of the proof

In this section, we outline the proof of Proposition 1.4. To begin with, fix w, one has by definition that

$$\Gamma_\delta(\phi_\delta(t)e^{it\Delta_\gamma}u_0+w)=i\phi_\delta(t)\int_0^t e^{i(t-s)\Delta_\gamma}\mathcal{N}(\phi(t)e^{it\Delta_\gamma}u_0+\phi(t)w)\,ds=:A+B, \tag{2.1}$$

where, according with (1.4), A corresponds to $\mathcal{N}_1$ and B to $\mathcal{N}_2$ respectively. We also use $\phi(t)\phi(t/\delta)=\phi(t)$. The estimate for part B follows from standard $X^{s,b}$ space estimates, which we present in the end of Sect. 3.2 for the sake of completeness. In order to study part A, we consider the Wick ordered nonlinearity $\mathcal{N}_1$ as a trilinear expression, replacing the w above by three functions w_1, w_2, w_3.

Using $X^{s,b}$ smoothing (3.8) and duality, Proposition 1.4 will follow from

Proposition 2.1 *There exist $\delta, \delta_0, \alpha_0, \epsilon_0, b$ as in Proposition 1.4 satisfying $\epsilon_0 \ll \epsilon_1 \ll s_0$, so that for any $h(x,t)$ with $\|h\|_{X^{0,1-b_0}} \le 1$, $h_0(x,t) := \phi(t/\delta)h(x,t)$, one has estimate*

$$\left|\int_{\mathbb{R}\times\mathbb{T}^2} <D>^{s_0}(\mathcal{N}_1(u_1,u_2,u_3))\overline{h_0}\,dxdt\right| \lesssim \delta^{\epsilon_1}, \tag{2.2}$$

$$\left|\int_{\mathbb{R}\times\mathbb{T}^2} <D>^{s_0}(\mathcal{N}_2(u_1,u_2,u_3))\overline{h_0}\,dxdt\right| \lesssim \delta^{\epsilon_1}, \tag{2.3}$$

where u_i is either $\phi_\delta(t)e^{it\Delta_\gamma}u_0$ or w_i.

Here $<\cdot>$ is the Japanese bracket, $<D>:=\sqrt{1-\Delta}$.

Remark 2.2 We will neglect any loss of $\delta^{-C\epsilon_0}$ throughout the proof, since eventually all such loss will be compensated by the gain of δ^{ϵ_1}. In particular, one should not be concerned about the loss in $X^{s,b}$ localization by multiplying $\phi(t/\delta)$.

In the two estimates above, (2.3) follows easily from deterministic estimates, whose proof will be given at the end of Sect. 3.2. The majority of the rest of the paper is devoted to proving (2.2).

More precisely, the proof of (2.2) splits into eight different cases depending on whether the input functions u_i are of the *regular* (in the space X^{s_0,b_0}) or *probabilisitic* forms. In addition, we further decompose each u_i into pieces corresponding to different spatial Fourier frequencies (i.e. replacing u_i with $P_{N_i}u_i$ for some dyadic numbers N_i), then the desired result follows from a case by case study depending on the relative sizes of the spatial frequencies N_1, N_2, N_3. Note that the roles of N_1 and N_3 are completely symmetric as shown in the definition of $\mathcal{N}$, so without loss of generality

we may assume $N_1 \geq N_3$ throughout. There are thus two main cases: $N_2 \geq N_1 \geq N_3$ and $N_1 \geq N_2$.

The first case $N_2 \geq N_1 \geq N_3$ turns out to be easier, which we treat in Sect. 5. The second case needs to be further decomposed depending on the relative sizes of N_2, N_3 and where the random terms appear. Following Bourgain's notation, we will use II to denote the regular case (i.e. $u_i = w_i$) and I to denote the probabilistic case (i.e. u_i is the cutoff of the free solution with random initial data u_0). In Sects. 6 and 7, we will first estimate two typical cases: $N_1(II) \geq N_2(I) \geq N_3(II)$ (corresponding to "case (a)" of [4]) and $N_1(I) \geq N_2(II) \geq N_3(II)$ (corresponding to "case (c)" of [4]). These two cases are typical in the sense that all essential elements of the proof and ideas will be displayed in the study of these two cases. Essentially this is because in the two cases, the random term appears in relatively higher frequencies hence there is less decay in terms of N_1 that one would expect; there is also no additional random term present which prevents one to fully exploit the cancellation brought by randomization. Note that these two sections are the main part of our proof. We will discuss the treatment of other cases in Sect. 8.

2.1 A list of cases

Following Bourgain, we need to study

- Case (0): $N_2 \gg N_1$;
- Case (a): $N_1(II) \geq N_2(I) \geq N_3(II)$;
- Case (b): $N_1(II) \geq N_3(I) \geq N_2(II)$;
- Case (c): $N_1(I) \geq N_2(II) \geq N_3(II)$;
- Case (d): $N_1(I) \geq N_3(II) \geq N_2(II)$;
- Case (e): $N_1(II) \geq N_2(I) \geq N_3(I)$;
- Case (f): $N_1(II) \geq N_3(I) \geq N_2(I)$;
- Case (g): $N_1(I) \geq N_2(I) \geq N_3(II)$;
- Case (h): $N_1(I) \geq N_3(I) \geq N_2(II)$;
- Case (i): $N_1(I) \geq N_2(II) \geq N_3(I)$;
- Case (j): $N_1(I) \geq N_3(I) \geq N_2(II)$;
- Case (k): $N_1(I) \geq N_2(I) \geq N_3(I)$;
- Case (l): $N_1(I) \geq N_3(I) \geq N_2(I)$.

Remark 2.3 Strictly speaking, one will need to study, for example in case (a), $N_2 \lesssim N_1$, $N_2 \geq N_3$. The analysis will be the same as for $N_1 \geq N_2 \geq N_3$, we neglect this issue.

2.2 Notation

For the sake of notational convenience, we will denote $\langle , \rangle_\gamma$ by $\langle , \rangle$ in short. We use P_N, $P_{\leq N}$ to denote Littlewood Paley projections in the physical space (x variable), as mentioned above. We will use $P_{|\tau|<M}$ as Littlewood Paley projections in the time space (t variable). We will also use $P_{|\tau-n^2|<M}$ to denote space time Littlewood Paley projections with respect to paraboloids.

For the sake of convenience, we sometimes abuse notation by identifying $P_N^2 = P_N$, $\phi(t)^2 = \phi(t)$. Throughout the paper, we use several parameters, and we always require

$$1 \geq s_0 \gg s_1 \geq \epsilon_1 \gg \epsilon_0 > 0. \tag{2.4}$$

3 Preliminaries

3.1 The $X^{s,b}$ space

In this subsection, we recall the definition of the $X^{s,b}$ space and summarize some classical estimates that will be used in the proof. One may refer to [2,5,6] for more details.

Let $v(t,x)$ be a function on $\mathbb{R} \times \mathbb{T}^2$. Let $\widehat{v}$ be the Fourier transform of v, i.e.

$$v(t,x) = \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}^2} \widehat{v}(n,\tau) e^{in\cdot x} e^{i\tau t}\, d\tau,$$

and the $X^{s,b}$ norm can be defined as

$$\|v\|_{X^{s,b}}^2 := \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}^2} < n >^{2s} < \tau - n^2 >^{2b} |\widehat{v}(n,\tau)|^2\, d\tau, \tag{3.1}$$

where $< n >:= \sqrt{1+n^2}$ is the Japanese bracket.

Note that another convenient way to define the $X^{s,b}$ norm is via the ansatz

$$v(t,x) = \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}^2} a(n,\lambda) e^{in\cdot x} e^{in^2 t} e^{i\lambda t}\, d\lambda, \tag{3.2}$$

which gives

$$\|v\|_{X^{s,b}}^2 = \int_{\mathbb{R}} \sum_{n \in \mathbb{Z}^2} |a(n,\lambda)|^2 < n >^{2s} < \lambda >^{2b}\, d\lambda. \tag{3.3}$$

The $X^{s,b}$ space is very useful in dispersive PDE for at least two reasons: first, it inherits the Strichartz estimates enjoyed by free solutions of the Schrödinger equation; second, it exploits the smoothing effect of the Duhamel formula.

We now recall the Strichartz estimates on tori, rational or irrational, [2,5],

$$\|e^{it\Delta_\gamma}(P_B f)\|_{L^4_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim_\epsilon N^\epsilon \|f\|_{L^2_x}, \tag{3.4}$$

where P_B is the Littlewood-Paley projection onto the spatial frequency ball B of radius N (not necessarily centered at the origin).

By the Minkowski inequality and Cauchy–Schwarz, this implies that

Lemma 3.1 *For any $u \in X^{0,\frac{1}{2}+\epsilon'}$, there holds*

$$\|P_B u\|_{L^4_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim_{\epsilon,\epsilon'} N^{\epsilon}\|u\|_{X^{0,\frac{1}{2}+\epsilon'}}. \tag{3.5}$$

Via an interpolation with the Hausdorff–Young inequality, the estimate above can be upgraded to

$$\|P_B u\|_{L^4_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim_{\epsilon} N^{\epsilon}\|u\|_{X^{0,\frac{1}{2}-\frac{\epsilon}{4}}}. \tag{3.6}$$

We also record another estimate, which follows immediately by interpolating (3.5) with the trivial bound

$$\|u\|_{L^2_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim \|u\|_{X^{0,0}}.$$

Lemma 3.2 *For any $u \in X^{0,\frac{1}{3}}$, there holds*

$$\|P_B u\|_{L^3_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim_{\epsilon} N^{\epsilon}\|u\|_{X^{0,\frac{1}{3}}}. \tag{3.7}$$

As mentioned earlier, the $X^{s,b}$ space also exploits the smoothing effect of the Duhamel formula, which can be made precise by the following estimate.

Lemma 3.3 *For all $s \geq 0, b > \frac{1}{2}$ and time cut-off function ϕ as above, there holds*

$$\left\|\phi(t)\int_0^t e^{i(t-s)\Delta_\gamma} v(s)\,ds\right\|_{X^{s,b}} \lesssim \|v\|_{X^{s,(b-1)}}. \tag{3.8}$$

Before ending this subsection, we also record the following localization properties of the $X^{s,b}$ space:

Lemma 3.4 *Let $u \in X^{s,b}$, then*

$$\|\phi_\delta(t)u\|_{X^{s,b}} \lesssim_b \begin{cases} \|u\|_{X^{s,b}}, & 0 < b < \frac{1}{2}, \\ \delta^{\frac{1}{2}-b}\|u\|_{X^{s,b}}, & \frac{1}{2} < b < 1. \end{cases} \tag{3.9}$$

Moreover, for all $0 \leq b' \leq b < 1/2$, there holds

$$\|\phi_\delta(t)u\|_{X^{s,b'}} \lesssim_{\epsilon} \delta^{b-b'-\epsilon}\|u\|_{X^{s,b}}. \tag{3.10}$$

3.2 Deterministic estimates

In this subsection, we collect several by now standard deterministic estimates. All of them were introduced when studying standard local theory of deterministic NLS on tori. We start with an estimate that exploits the time localization. One may refer to [2,6] for proof. We provide a brief sketch of proof of the lemma in Appendix A for the convenience of the reader.

Lemma 3.5 *Let P_B be the Littlewood-Paley projection onto the spatial frequency ball B of radius N and $0 < s \ll 1$. Then one has for all $\epsilon \ll s$ that*

$$\|\phi(t) P_B u\|_{L^4_{t,x}} \lesssim_\epsilon N^{C\epsilon} \|u\|_{X^{0,1/2-\epsilon}}, \tag{3.11}$$

$$\|\phi_\delta(t)(P_B u)\|_{L^4_{t,x}} \lesssim_\epsilon N^{s+C\epsilon} \|u\|_{X^{0,\frac{1}{2}-\epsilon}} \delta^{\frac{s}{8}}. \tag{3.12}$$

The number $1/8$ is not meant to be sharp, one can for example upgrade it to $1/4-$. Throughout the rest of the section, we write

$$f_i(x,t) = \sum_{n\in\mathbb{Z}^2} f_i(n,t) e^{in\cdot x}, \quad i = 1,2,3, \tag{3.13}$$

i.e. $f_i(n,t)$ is the space Fourier transform of f_i. For the sake of brevity, we abbreviate $f_i(n,t)$ as $f_i(n)$.

We summarize below several standard estimates that will be frequently used in the proofs that will come later. One may refer to [2,4,6]. We will also sketch them in Appendix B for the convenience of the reader.

Lemma 3.6 *Let $N_1 \gtrsim N_2, N_3$, $1 \gg s_1 \gg \epsilon_0$, and $\psi(t)$ be a Schwartz function. Decompose $P_{N_1} = \sum_{J\in\mathcal{J}} P_J$, where $J \in \mathcal{J}$ are finitely overlapping balls in the region $|n| \sim N_1$ with radius $\sim \max(N_2, N_3)$. then one has*

$$\left|\int \psi(t)\bar{h} P_{N_1} f_1 \overline{P_{N_2} f_2} P_{N_3} f_3\right| \lesssim (\max(N_2,N_3))^{C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}}, \tag{3.14}$$

$$\left|\int \psi(t)\bar{h}\mathcal{N}_1(P_{N_1} f_1, P_{N_2} f_2, P_{N_3} f_3)\right| \lesssim (\max(N_2,N_3))^{C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}}, \tag{3.15}$$

$$\left|\int \psi(t/\delta)\bar{h} P_{N_1} f_1 \overline{P_{N_2} f_2} P_{N_3} f_3\right| \lesssim \delta^{s_1/8}(\max(N_2,N_3))^{s_1+C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}}, \tag{3.16}$$

$$\left|\int \psi(t/\delta)\bar{h}\mathcal{N}_1(P_{N_1} f_1, P_{N_2} f_2, P_{N_3} f_3)\right| \lesssim \delta^{s_1/8}(\max(N_2,N_3))^{s_1+C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}}, \tag{3.17}$$

$$\begin{aligned}&\left|\int \psi(t)\bar{h}\mathcal{N}_1(P_{N_1} f_1, P_{N_2} f_2, P_{N_3} f_3)\right| \\ &\lesssim (\max(N_2,N_3))^{C\epsilon_0} \|h\|_{X^{0,1/3}} \|f_2\|_{X^{0,1/3}} \|f_3\|_{X^{0,1/3}} \sup_J \|P_J f_1\|_{L^\infty_{t,x}} \\ &\quad + \mathbf{1}_{N_1\sim N_2} \|P_{N_1} f_1\|_{X^{0,b_0}} \|P_{N_1} f_2\|_{X^{0,1/3}} \|P_{N_3} f_3\|_{X^{0,1/3}} \|P_{N_3} h\|_{X^{0,1/3}}\end{aligned}$$

$$+\mathbf{1}_{N_2\sim N_3}\|P_{N_1}f_1\|_{X^{0,b_0}}\|P_{N_2}f_2\|_{X^{0,1/3}}\|P_{N_2}f_3\|_{X^{0,1/3}}\|P_{N_1}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2\sim N_3}\min_i\left(\|P_{N_1}h\|_{X^{0,1-b_0}}\|f_i\|_{X^{0,b_0}}\sup_{|n|\sim N_1}\prod_{j\neq i}\|f_j(n)e^{in\cdot x}\|_{X^{0,b_0}}\right), \tag{3.18}$$

$$\left|\int\psi(t)\bar{h}\mathcal{N}_1(P_{N_1}f_1,P_{N_2}f_2,P_{N_3}f_3)\right|$$
$$\lesssim N_1^{C\epsilon_0}\|h\|_{X^{0,1/3}}\|P_{N_1}f_1\|_{L^\infty_{t,x}}\|f_2\|_{X^{0,1/3}}\|f_3\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2}\|P_{N_1}f_1\|_{X^{0,b_0}}\|P_{N_1}f_2\|_{X^{0,1/3}}\|P_{N_3}f_3\|_{X^{0,1/3}}\|P_{N_3}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_2\sim N_3}\|P_{N_1}f_1\|_{X^{0,b_0}}\|P_{N_2}f_2\|_{X^{0,1/3}}\|P_{N_2}f_3\|_{X^{0,1/3}}\|P_{N_1}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2\sim N_3}\min_i\left(\|P_{N_1}h\|_{X^{0,1-b_0}}\|f_i\|_{X^{0,b_0}}\sup_{|n|\sim N_1}\prod_{j\neq i}\|f_j(n)e^{in\cdot x}\|_{X^{0,b_0}}\right), \tag{3.19}$$

$$\left|\int\psi(t)\bar{h}\mathcal{N}_1(P_{N_1}f_1,P_{N_2}f_2,P_{N_3}f_3)\right|$$
$$\lesssim(\max(N_2,N_3))^{C\epsilon_0}\|h\|_{X^{0,1/3}}\|P_{N_2}f_2\|_{L^\infty_{t,x}}\|P_{N_1}f_1\|_{X^{0,1/3}}\|P_{N_3}f_3\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2}\|P_{N_1}f_1\|_{X^{0,1/3}}\|P_{N_1}f_2\|_{X^{0,b_0}}\|P_{N_3}f_3\|_{X^{0,1/3}}\|P_{N_3}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_2\sim N_3}\|P_{N_1}f\|_{X^{0,1/3}}\|P_{N_2}f_2\|_{X^{0,b_0}}\|P_{N_2}f_3\|_{X^{0,1/3}}\|P_{N_1}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2\sim N_3}\min_i\left(\|P_{N_1}h\|_{X^{0,1-b_0}}\|f_i\|_{X^{0,b_0}}\sup_{|n|\sim N_1}\prod_{j\neq i}\|f_j(n)e^{in\cdot x}\|_{X^{0,b_0}}\right), \tag{3.20}$$

$$\left|\int\psi(t)\bar{h}\mathcal{N}_1(P_{N_1}f_1,P_{N_2}f_2,P_{N_3}f_3)\right|$$
$$\lesssim(\max(N_2,N_3))^{C\epsilon_0}\|h\|_{X^{0,1/3}}\|P_{N_2}f_2\|_{X^{0,1/3}}\|P_{N_1}f_1\|_{X^{0,1/3}}\|P_{N_3}f_3\|_{L^\infty_{t,x}}$$
$$+\mathbf{1}_{N_1\sim N_2}\|P_{N_1}f_1\|_{X^{0,1/3}}\|P_{N_1}f_2\|_{X^{0,1/3}}\|P_{N_3}f_3\|_{X^{0,b_0}}\|P_{N_3}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_2\sim N_3}\|P_{N_1}f\|_{X^{0,1/3}}\|P_{N_2}f_2\|_{X^{0,1/3}}\|P_{N_2}f_3\|_{X^{0,b_0}}\|P_{N_1}h\|_{X^{0,1/3}}$$
$$+\mathbf{1}_{N_1\sim N_2\sim N_3}\min_i\left(\|P_{N_1}h\|_{X^{0,1-b_0}}\|f_i\|_{X^{0,b_0}}\sup_{|n|\sim N_1}\prod_{j\neq i}\|f_j(n)e^{in\cdot x}\|_{X^{0,b_0}}\right), \tag{3.21}$$

where with $\mathbf{1}_{N_i\sim N_j}$ *we denote the indicator* $\mathbf{1}_{N_i\sim N_j}=1$ *if* $N_i\sim N_j$, 0 *otherwise. Moreover, estimate* (3.18), (3.19), (3.20), (3.21) *are also valid if one replaces* $\mathcal{N}_1(P_{N_1}f_1,P_{N_2}f_2,P_{N_3}f_3)$ *by* $P_{N_1}f_1\overline{P_{N_2}f_2}P_{N_3}f_3$.

Similarly, one also has

Lemma 3.7 *If $N_2 \gtrsim N_1 \geq N_3$, one has*

$$\begin{aligned}&\left|\int \phi(t/\delta)\bar{h} P_{N_1} f_1 \overline{P_{N_2} f_2} P_{N_3} f_3\right| \\ &\lesssim \delta^{s_1/8} N_1^{s_1} N_1^{C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}}, \end{aligned} \tag{3.22}$$

$$\begin{aligned}&\left|\int \phi(t)\bar{h} P_{N_1} f_1 \overline{P_{N_2} f_2} P_{N_3} f_3\right| \\ &\lesssim N_1^{C\epsilon_0} \|h\|_{X^{0,1/3}} \|P_{N_i} f_i\|_{L^\infty_{t,x}} \prod_{j\neq i} \|P_{N_i} f_i\|_{X^{0,1/3}}, \quad i = 1, 2, 3. \end{aligned} \tag{3.23}$$

We also record the following deterministic estimate which will (almost directly) handle the $\mathcal{N}_2$ part in the Wick ordered nonlinearity.

Lemma 3.8 *Let $b_0 = \frac{1}{2} + \epsilon_0$, then*

$$\begin{aligned}&\left|\int \phi(t/\delta)\bar{h}\mathcal{N}_2(P_{N_1} f_1, P_{N_1} f_2, P_{N_1} f_3)\right| \\ &\lesssim \min_i \left(\delta^{1/4} \|P_{N_1} h\|_{X^{0,1-b_0}} \|f_i\|_{X^{0,b_0}} \sup_{|n|\sim N_1} \prod_{j\neq i} \|f_j(n) e^{in\cdot x}\|_{X^{0,b_0}}\right). \end{aligned} \tag{3.24}$$

We sketch the proofs of Lemma 3.6, 3.7 and 3.8 in Appendix B for the convenience of the reader. In the following, we provide a proof of the easier estimate in our main result Proposition 2.1.

Proof of (2.3) of Proposition 2.1 We choose N_0 large, up to dropping a set of probability $e^{-N_0^{cs_1}}$, we have

$$\begin{cases} |g_n(\omega)| \leq N_0^{s_1}, |n| \leq N_0, \\ |g_n(\omega)| \leq |n|^{s_1}, |n| \geq N_0. \end{cases} \tag{3.25}$$

And in particular, no matter whether $u_i = w_i$ such that $\|w_i\|_{X^{s_0,b_0}} \lesssim 1$ or $u_i = \phi(t) \sum_{|n|\sim N_i} \frac{g_n(\omega)}{|n|} e^{in^2 t} e^{inx}$, we always have

$$\begin{cases} \|P_N u_i|_{X^{0,b_0}} \lesssim N_0^{s_1}, N \leq N_0, \\ \|P_N u_i\|_{X^{0,b_0}} \lesssim N^{s_1}, |N| \geq N_0, \end{cases} \tag{3.26}$$

Observe that

$$\int \phi(t/\delta)\bar{h}\mathcal{N}_2(u_1, u_2, u_3) = \sum_N \int \phi(t/\delta) P_N \bar{h}\mathcal{N}_2(P_N u_1, P_N u_2, P_N u_3) \tag{3.27}$$

Let $\delta = N_0^{-100}$, applying estimate (3.24), we have

- $N \le N_0$
$$\int \phi(t/\delta) P_N \bar{h} \mathcal{N}_2(P_N u_1, P_N u_2, P_N u_3) \lesssim \delta^{1/4} N^{Cs_1}, \tag{3.28}$$
- $N \ge N_0$,
$$\int \phi(t/\delta) P_N \bar{h} \mathcal{N}_2(P_N u_1, P_N u_2, P_N u_3) \lesssim N^{Cs_1} N^{-2}, \tag{3.29}$$

Sum all N, (when s_0 small enough), and desired estimate follow. □

3.3 Probabilistic estimates

We collect the elementary but crucial probabilistic estimates here.

Lemma 3.9 *Let $\{g_n(\omega)\}$ be i.i.d complex Gaussian on the probability space Ω, and $\{c_{n_1,\dots,n_k}\}$ be a sequence of complex numbers for some integer $k \ge 1$. Define*

$$F_k(\omega) = \sum_{n_1,\dots,n_k} c_{n_1,\dots,n_k} g_{n_1} g_{n_2} \cdots g_{n_k}.$$

Then one has for all $1 < p < \infty$

$$\|F_k\|_{L^p(\Omega)} \lesssim \sqrt{k+1}(p-1)^{k/2} \|F_k\|_{L^2(\Omega)}. \tag{3.30}$$

Moreover, there holds the associated large deviation type estimate

$$\mathbb{P}\{|F_k| > \lambda\} \le \exp\left(\frac{-C\lambda^{2/k}}{\|F_k\|^{2/k}_{L^2(\Omega)}}\right), \qquad \forall \lambda > 0. \tag{3.31}$$

In the lemma above, it is very important that $\{c_{n_1,\dots,n_k}\}$ are numbers instead of random variables. One may refer to [11,12].

The following lemma will also be frequently used.

Lemma 3.10 *Let $\{g_n(\omega)\}$ be i.i.d complex Gaussian on the probability space Ω, and assume that*

$$\sum_{n\in\mathbb{Z}^2} |a_n|^2 \lesssim 1.$$

Then, for any integer $N > 0$, up to a set of probability measure e^{-N^α} for some $\alpha > 0$ depending on ϵ, there holds

$$\left\| \sum_{n\in\mathbb{Z}^2, |n|\le N} a_n g_n(\omega) e^{in\cdot x} \right\|_{L^\infty(\mathbb{T}^2)} \lesssim N^\epsilon, \quad \forall \epsilon > 0. \tag{3.32}$$

Proof It is easy to see that for any fixed x, the function is bounded as desired outside a small exceptional set, so the key point is to show that the exceptional set can be made independent of x. To do this, given $\epsilon > 0$, first note that $\mathbb{T}^2$ can be covered by a mesh of size $1/N^M \times 1/N^M$ centered at $\sim N^{2M}$ lattice points for a large number M to be determined later. We first bound the function at the lattice points, which is easy as the function at each lattice point has size $\lesssim N^{\epsilon_0}$ (for some $\epsilon_0 < \epsilon$) up to probability $e^{-N^{\alpha(\epsilon_0)}}$ according to Lemma 3.9, and there are only N^{2M} many points. Therefore, one has that the function satisfies the desired bound outside an exceptional set of measure up to $\sim e^{-N^{\alpha-}}$.

To pass from here to the bound of the function on the entire $\mathbb{T}^2$, it suffices to obtain a uniform control of the derivative of the function (independent of x):

$$N \sum_{|n|\leq N} |a_n||g_n(\omega)| \lesssim N^2 \sup_n |g_n(\omega)|. \tag{3.33}$$

The probability of the derivative being larger than N^4 is smaller than $e^{-N^2}N^2$, as the probability of $|g_n(\omega)| > N^2$ for each n is controlled by e^{-N^2}. Hence, by removing this additional exceptional set and recalling that every point x lies within the distance of $1/N^M$ from some lattice point, one has that the function at x is bounded by $N^{\epsilon_0} + N^4 \cdot 1/N^M \lesssim N^{\epsilon}$ as long as M is chosen sufficiently large. □

4 Counting lemma

One of the key ingredients in the proof of our main theorem is an extension of the lattice counting argument of Bourgain, [4] to the irrational setting. We present them in this section. We start with two auxiliary lemmata. The first has a geometric flavor, while the second is an elementary number theoretical result.

Lemma 4.1 *Let A be the $O(\frac{1}{N})$-neighborhood of a circle of radius $\sim N$, and $\Lambda = \mathbb{Z} \times \gamma\mathbb{Z}$ for some real number $\gamma \in (1, 2)$. Suppose $A_1 \subset A$ is the $O(\frac{1}{N})$ neighborhood of an arc of the circle of length $N_2 \leq N$. Then, A_1 contains at most* $\max\left(\frac{N_2}{N^{1/3}}, 1\right)$ *points of* Λ.

Proof Let $C \subset A$ be the $O(\frac{1}{N})$-neighborhood of any arc of the circle of angular size $\theta = \frac{1}{1000}N^{-2/3}$, then it suffices to show that C contains at most $O(1)$ points from Λ. Indeed, if $N_2 > N^{1/3}$, A_1 corresponds to an arc of angular size $\frac{N_2}{N}$, which can be decomposed into $\sim \frac{N_2}{N^{1/3}}$ smaller arcs each of which containing at most $O(1)$ points from Λ.

Denote B_1 the circular sector bounded by the outer arc of C, and B_2 the triangle with vertices being the center of the circle and the two endpoints of the inner arc of C. Observe that any triangle $P_1P_2P_3$ with $P_i \in \Lambda \cap C$ must be contained in the region $B_1 - B_2$. Moreover, it is easy to see that annulus A can contain straight line of length at most $O(1)$. Therefore, suppose C contains more than $O(1)$ points from Λ, then there must exist three points $P_1, P_2, P_3 \in \Lambda \cap C$ that formulate a non-degenerate triangle.

By definition of Λ, the area of the triangle is at least $\frac{1}{2}$, hence the area of $B_1 - B_2$ needs to be at least $\frac{1}{2}$ as well.

On the other hand, via Taylor expansion, the area of $B_1 - B_2$ is bounded by $\frac{1}{2}N^2(\theta - \sin\theta) + O(\frac{1}{N})N\theta \lesssim N^2\theta^3 + \theta \leq \frac{1}{10}$, which is a contradiction. Therefore, C must contain at most $O(1)$ points from Λ and the proof is complete. □

Lemma 4.2 *Given an integer $M \neq 0$, then*

$$\#\{(a,b) \in \mathbb{Z} \times \mathbb{Z} : ab = M\} \leq C_\epsilon |M|^\epsilon, \quad \forall \epsilon > 0.$$

Proof Without loss of generality we assume that $M > 0$. As an integer, M has an unique representation by its prime factors:

$$M = p_1^{r_1} p_2^{r_2} \cdots p_\ell^{r_\ell}, \quad p_1 < p_2 < \cdots < p_\ell, \quad r_i > 0, \ \forall i.$$

Then, the total number of pairs of integers whose product is M is bounded by $\prod_{i=1}^{\ell}(r_i + 1)$. For any fixed $\epsilon > 0$, there exists a smallest integer N such that $N^\epsilon > 2$. Let p_j be the first factor that is larger than N, then there holds

$$\prod_{i=j}^{\ell}(r_i + 1) \leq \prod_{i=j}^{\ell} 2^{r_i} < \prod_{i=j}^{\ell} p_i^{r_i \epsilon} < M^\epsilon.$$

On the other hand, there are only $O_\epsilon(1)$ many p_i that are smaller than N. Therefore, write $M = e^m$, one has

$$\prod_{i=1}^{j-1}(r_i + 1) \leq (\log M)^{O_\epsilon(1)} = m^{O_\epsilon(1)}.$$

There exists a large number $M_0 = M_0(\epsilon)$ so that $m^{O_\epsilon(1)} \leq e^{\epsilon m}$ whenever $M = e^m > M_0$, hence the desired estimate follows if $M > M_0$. If $M \leq M_0$, one can simply take $C_\epsilon = M_0^2$. The proof is complete. □

We now fix μ and $N_1, N_2, N_3 \geq 0$, and we let

$$S := \{(n_1, n_2, n_3) : |n_i| \sim N_i, \ n_2 \neq n_1, n_3, \text{ and } \langle n_2 - n_1, n_2 - n_3\rangle = \mu + O(1)\}.$$

We observe here that in the rational case S is a curve while in the general case, since $+O(1)$ appears the set is thick.

Define

$$S(n_1) = \{(n_2, n_3) : (n_1, n_2, n_3) \in S\},$$

and similarly for other $S(n_i)$, if n_i is fixed and $S(n_i, n_j)$, if n_i, n_j are fixed. We have the following counting lemmata regarding the size of these sets. In the following, we

sometimes use N^1, N^2, N^3 to denote N_1, N_2, N_3 rearranged in the non-increasing order and assume $\mu = O(N^1)$.

Lemma 4.3 *$\#S(n_1, n_2) \lesssim N_3$ and $\#S(n_2, n_3) \lesssim N_1$.*

Compared to the rational case studied by Bourgain, this estimate is equally good, since ultimately it is a linear estimate.

Proof We will only prove the first estimate, as the second one follows from the same argument. Fixing $n_1 \neq n_2$, one has from

$$\langle n_2 - n_1, n_2 - n_3 \rangle = \mu + O(1)$$

that n_3 lies in an $O(\frac{1}{|n_1 - n_2|}) \leq O(1)$ neighborhood of a straight line. Since $|n_3| \sim N_3$, there are at most $\sim N_3$ choices of n_3. □

It is in the next lemmata that one sees a difference with respect to the estimates of Bourgain that are generated by the possible irrationality of the torus.

Lemma 4.4 *Assume $N_1 \geq N_2, N_3$. Then,*

$$\#S(n_1, n_3) \lesssim \begin{cases} N_2^{2/3}, & \text{if } N_1 \sim N_3 \gg N_2, \\ \max\left(\frac{N_2}{(N_1)^{1/3}}, 1\right), & \text{otherwise.} \end{cases}$$

Proof From the definition of the set S, with n_1, n_3 fixed, n_2 must lie in an annulus given by the formula

$$\left| n_2 - \frac{n_1 + n_3}{2} \right|^2 = \frac{|n_1 - n_3|^2}{4} + \mu + O(1).$$

Denote the inner and outer radius of the annulus by R_1, R_2 respectively and recall that $\mu \lesssim N_1$.

Therefore, when $N_1 \gg N_3$, both the inner and outer radius are roughly $\sim \sqrt{(N_1)^2 + \mu} \sim N_1$. In order to determine the thickness of the annulus, one observes from $R_1^2 - R_2^2 = O(1)$ and $R_1, R_2 \sim N_1$ that there holds $R_1 - R_2 \leq O(\frac{1}{N_1})$, hence the thickness is bounded by $O(\frac{1}{N_1})$. Then, the desired estimate $\max(\frac{N_2}{N_1^{1/3}}, 1)$ follows immediately from $|n_2| \sim N_2 \leq N_1$ and Lemma 4.1 above.

When $N_1 \sim N_3$, assume that the inner and outer radius are roughly $\sim R \gg 1$ (if $R \leq O(1)$, the estimate is trivial). Note that $R \lesssim N_1$. Then n_2 lies inside an annulus of radius $\sim R$ and thickness bounded by $O(\frac{1}{R})$. Suppose $N_1 \sim N_2 \sim N_3$, then again by Lemma 4.1 above, the total number of n_2 is bounded by

$$\max\left(\frac{R}{R^{1/3}}, 1\right) = \max(R^{2/3}, 1) \lesssim \max(N_1^{2/3}, 1) \sim \max\left(\frac{N_2}{N_1^{1/3}}, 1\right).$$

On the other hand, if $N_1 \sim N_3 \gg N_2$, still denoting R as roughly the inner and outer radius of the annulus, one has

$$\#S(n_1, n_3) \lesssim \max\left(\frac{\min(N_2, R)}{R^{1/3}}, 1\right) \lesssim N_2^{2/3}.$$

□

Lemma 4.4 above can be extended to estimate other sets of similar type. For example, let $n := n_1 - n_2 + n_3$ and suppose that $N_1 = N^1$. Then for any fixed n_2 and n, via a similar argument, one has the following estimate:

$$\#\{n_1 : |n_i| \sim N_i,\ n_2 \neq n_1, n_3, \text{ and } \langle n_1 - n, n_2 - n_1\rangle = \mu + O(1)\} \lesssim N_1^{2/3}.$$

Indeed, suppose $N_1 \gg N_2$, then n_1 in the above lies in an annulus of radius $\sim N_1$ and thickness $\sim O(\frac{1}{N_1})$. Hence by Lemma 4.1, the total number is at most $\max(\frac{N_1}{N_1^{1/3}}, 1) = N_1^{2/3}$. Otherwise, if $N_1 \sim N_2$, one has R, the radius of the annulus, is bounded by N_1. Hence, the total possible number of n_1 is at most $\lesssim R^{2/3} \lesssim N_1^{2/3}$.

Similarly, when $N_1 \gg N_2$, one also has the following counting

$$\#\{n_3 : \langle n_3 - n_2, n_3 - n\rangle = \mu + O(1)\} \lesssim \max\left(\frac{N_3}{N_1^{1/3}}, 1\right).$$

Moreover, from the two lemmata above, one can already obtain some estimate for sets $S(n_i)$. For instance, by first fixing n_2 and applying Lemma 4.3, one can show that $\#S(n_1) \lesssim N_2^2 N_3$. Depending on the relative sizes the N_i, sometimes such estimates are already good enough. However, in some other cases one needs to use a more sophisticated argument, and this is the contents of the following counting lemma.

Lemma 4.5 *$\#S(n_1) \lesssim (N^1)^\epsilon N_2 N_3$, $\#S(n_2) \lesssim (N^1)^\epsilon N_1 N_3$, and $\#S(n_3) \lesssim (N^1)^\epsilon N_1 N_2$.*

Proof We only prove the estimate of $\#S(n_1)$ as the other two can be treated very similarly. Write $n_2 - n_1 = r(a, \gamma b)$, where $r \in \mathbb{N}$, $a, b \in \mathbb{Z}$ and $(a, b) = 1$. Decompose all choices of n_2 into dyadic scales. In other words, at each scale, we have dyadic number $A, B \in \mathbb{Z}$ fixed such that $|a| \sim A$, $|b| \sim B$, and there holds $A, B \lesssim \max(N_1, N_2)$. We also write $n_2 - n_3 = (x, \gamma y)$, $x, y \in \mathbb{Z}$.

Assume $a, b \neq 0$ and fix A, B, r. We want to count the number of (a, b, x, y) satisfying

$$r(ax + \gamma^2 by) = \mu + O(1), \quad \text{and } |a| \sim A,\ |b| \sim B.$$

Note that $r \neq 0$ because $n_1 \neq n_2$, and x, y cannot both be zero as $n_2 \neq n_3$. Without loss of generality, suppose $y \neq 0$, then the equality above can be rewritten as

$$\frac{ax}{by} + \gamma^2 = \frac{\mu}{byr} + O\left(\frac{1}{|byr|}\right).$$

Since $a, b, x, y \in \mathbb{Z}$, for any fixed value $H = by$, the value of ax is inside an $O(1)$-neighborhood of an integer $G = G(H)$.

Moreover, observe that the number of possible values of H is bounded by $\sim BN_3$, as for each fixed b (hence the second coordinate of n_2 is fixed) there are $\sim N_3$ many choices of y. Then, by a simple number theory observation (Lemma 4.2 below) one has for any $\epsilon > 0$ that

$$\begin{aligned}\#(a, b, x, y) &\lesssim \#\{H\} \cdot |H|^{\epsilon} \cdot |G|^{\epsilon} \lesssim BN_3(B\max(N_2, N_3))^{\epsilon}(A\max(N_2, N_3))^{\epsilon} \\ &\lesssim \max(A, B)^{1+\epsilon} N_3 \max(N_2, N_3)^{\epsilon}.\end{aligned}$$

It is thus left to sum over r and then A, B. Note that for fixed A, B,

$$\#r \lesssim \frac{N_2}{\max(A, B)},$$

therefore, one has in this case that

$$\begin{aligned}\#S(n_1) &\lesssim \sum_{A,B} \frac{N_2}{\max(A, B)} \max(A, B)^{1+\epsilon} N_3 \max(N_2, N_3)^{\epsilon} \\ &\lesssim \log(N^1)^2 (N^1)^{\epsilon} N_2 N_3 \lesssim (N^1)^{\epsilon} N_2 N_3.\end{aligned}$$

Assume now that $a = 0$ (then $b \neq 0$ as $n_2 \neq n_1$). This means n_1, n_2 have the same first coordinate, hence the total number of choices of n_2 is bounded by N_2. Moreover, one has that the first coordinate of n_3 is free and its second coordinate is determined by

$$\langle n_2 - n_1, n_2 - n_3 \rangle = \mu + O(1),$$

hence is inside an $O(1)$-neighborhood of a determined value. Indeed, the formula above can be written as

$$r\gamma^2 by = \mu + O(1)$$

which implies

$$y = \frac{\mu}{r\gamma^2 b} + O\left(\frac{1}{|r\gamma^2 b|}\right).$$

Therefore, in this case one has $\#S(n_1) \lesssim N_2 N_3$. The $b = 0$ case can be treated in the same way which we omit. □

5 Proof of Proposition 2.1: case by case study, case 0

In this section, we treat the case: $N_2 \gg N_1 \geq N_3$. We will prove

$$\left| \sum_{N_2 \gg N_1 \geq N_3} N_2^{s_0} \int \mathcal{N}_1(P_{N_1}u_1, P_{N_2}u_2, P_{N_3}u_3)\bar{h}\phi(t/\delta) \right| \lesssim \delta^{\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0, \tag{5.1}$$

where u_i is either w_i with $\|w_i\|_{X^{s_0,b_0}} \leq 1$ or

$$u_i = \phi(t)e^{it\Delta_\gamma}u_0 = \phi(t)e^{it\Delta_\gamma}\sum_n \frac{g_n(\omega)}{|n|}e^{in\cdot x}, \quad x \in \mathbb{T}^2,$$

and $\|h\|_{X^{0,1-b_0}} \leq 1$.

First observe that, in this case, the Wick ordered nonlinearity is the same as the usual cubic nonlinearity, i.e. $\mathcal{N} = \mathcal{N}_1$ and $\mathcal{N}_2 = 0$. We only need to prove (up to an exceptional set of probability $e^{-\delta^{-c}}$) that

$$\left| \sum_{N_2 \gg N_1 \geq N_3} N_2^{s_0} \int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1}u_1\overline{P_{N_2}u_2}P_{N_3}u_3\bar{h}\phi(t/\delta) \right| \lesssim \delta^{\epsilon_1} \text{ for some } \epsilon_1 \gg \epsilon_0. \tag{5.2}$$

There are several subcases. We start with **subcase 1** : $u_i = w_i$, $i = 1, 2, 3$. Let s_1 be chosen such that $\epsilon_0 \ll s_1 \ll s_0$. Via (3.22) and observing that one has $h = P_{N_2}h$ in the being in this case, we obtain

$$\begin{aligned} &\left| \int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1}u_1\overline{P_{N_2}u_2}P_{N_3}u_3\bar{h}\phi(t/\delta) \right| \\ &\lesssim N_1^{2s_1}\|P_{N_1}u_1\|_{X^{0,b_0}}\|P_{N_2}u_2\|_{X^{0,b_0}}\|P_{N_3}u_3\|_{X^{0,b_0}}\|P_{N_2}h\|_{X^{0,1-b_0}}\delta^{\frac{s_1}{8}} \\ &\lesssim N_1^{-s_0/2}\|u_1\|_{X^{s_0,b_0}}\|u_3\|_{X^{s_0,b_0}}\|P_{N_2}u_2\|_{X^{0,b_0}}\|P_{N_2}h\|_{X^{0,b_0}}. \end{aligned} \tag{5.3}$$

Sum over $N_2 \gg N_1 \geq N_3$, and the desired estimate follows.

Next we discuss **subcase 2**: at least one u_i is $\phi(t)\sum_{|n_i|\sim N_i}\frac{g_{n_i}}{|n_i|}e^{in_ix}e^{in_i^2t}$. We only study the case $u_1 = \phi(t)\sum_{|n_1|\sim N_1}\frac{g_{n_1}}{|n_1|}e^{in_1x}e^{in_1^2t}$, as other cases can be treated similarly.

Let $N_{2,0}$ be a large parameter such that $N_{2,0}^{100} = \frac{1}{\delta}$. Note that up to an exceptional set of probability $e^{-\delta^{-c}} \sim e^{-N_{2,0}^{c'}}$, we have

$$\begin{cases} |g_n(\omega)| \leq N_{2,0}^{s_1}, & |n| \leq N_{2,0}, \\ |g_n(\omega)| \leq N_2^{s_1}, & |n| \sim N_2 \geq N_{2,0}. \end{cases} \tag{5.4}$$

In particular, one always has for all u_i that

$$\begin{cases} \|P_N u_i\|_{X^{0,b_0}} \lesssim N_{2,0}^{s_1}, & N \le N_{2,0}, \\ \|P_N u_i\|_{X^{0,b_0}} \lesssim N^{s_1}, & N \ge N_{2,0}. \end{cases} \tag{5.5}$$

and, dropping another exceptional set of probability $e^{-\delta^{-c}}$ if necessary, one has

$$\|P_{N_1} u_1\|_{L^\infty_{t,x}} \le N_1^{s_1}, \quad N_1 \le N_2, \quad N_{2,0} \le N_2. \tag{5.6}$$

Now we split the **subcase 2** further into the following subsubcases.

In **subsubcase 2.1**, we restrict ourselves to the regime $N_2 \le N_{2,0}$, and use estimate (3.22) to derive

$$\begin{aligned} &\left| \int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1} u_1 \overline{P_{N_2} u_2} u_3 \bar{h} \phi(t/\delta) \right| \\ &\lesssim \delta^{s_1/8} N_1^{2s_1} \|u_1\|_{X^{0,b_0}} \|u_2\|_{X^{0,b_0}} \|u_3\|_{X^{0,b_0}} \|h\|_{X^{0,1-b_0}} \lesssim \delta^{s_1} N_1^{2s_1}. \end{aligned} \tag{5.7}$$

Summing over $N_1, N_3 \le N_2 \le N_{2,0}$, one obtains $\lesssim \delta^{s_1} N_{2,0}^{3s_1}$ and the desired estimate follows.

We are left with **subsubcase 2.2**, where $N_2 \ge N_{2,0}$. We will prove

$$\left| \int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1} \phi(t) u_1 \overline{P_{N_2} \phi(t) u_2} \phi(t) P_{N_3} u_3 \phi(t) \bar{h} \right| \lesssim N_2^{-1/10}, \tag{5.8}$$

which will then imply the desired estimate by summing (5.8) over N_1, N_2, N_3.

As remarked above, we only prove estimate (5.8) for the case that u_1 is random. Using (3.23) and Lemma 3.10, one derives

$$\begin{aligned} &\left| \int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1} \phi(t) u_1 \overline{P_{N_2} \phi(t) u_2} \phi(t) P_{N_3} u_3 \phi(t) \bar{h} \right| \\ &\lesssim N_2^{2s_1} \|\phi(t) P_{N_2} u_2\|_{X^{0,1/3}} \|\phi(t) P_{N_3} u_3\|_{X^{0,1/3}} \|\phi(t) P_{N_2} h\|_{X^{0,1/3}}. \end{aligned} \tag{5.9}$$

Let $F_i(\tau_i, n_i)$ be the space-time Fourier transform of $\phi(t)u_i$, $i = 1, 2, 3$, and $F_4(\tau_4, n_4)$ be the Fourier transform of $\phi(t)h$, the integral being estimated is non-zero only if

$$\begin{cases} n_1 - n_2 + n_3 - n_4 = 0, \\ \tau_1 - \tau_2 + \tau_3 - \tau_4 = 0, \end{cases} \tag{5.10}$$

which implies

$$\sum_i (-1)^i (\tau_i - n_i^2) = n_1^2 - n_2^2 + n_3^2 - n_4^2 = -2\langle n_2 - n_1, n_2 - n_3 \rangle \sim N_2^2, \text{ since } N_2 \gg N_1, N_3. \tag{5.11}$$

Observe that the Fourier transform of $\phi(t)u_1$ is essentially supported on $|\tau_1 - n_1^2| \lesssim 1$, thus, at least for one $i \in \{2, 3, 4\}$, one has $N_2^2 \lesssim |\tau_i - n_i^2|$, hence one can upgrade estimate (5.9) to

$$\begin{aligned}&\left|\int_{\mathbb{R}\times\mathbb{T}^2} P_{N_1}\phi(t)u_1\overline{P_{N_2}\phi(t)u_2}\phi(t)P_{N_3}u_3\phi(t)\bar{h}\right|\\&\lesssim N_2^{2s_1}N_2^{2(-1/6-\epsilon_0)}\|\phi(t)P_{N_2}u_2\|_{X^{0,b_0}}\|\phi(t)P_{N_3}u_3\|_{X^{0,b_0}}\|\phi(t)P_{N_2}h\|_{X^{0,1-b_0}}\\&\lesssim N_2^{Cs_1}N_2^{-1/6}\lesssim N_2^{-1/10}.\end{aligned} \tag{5.12}$$

To make the above argument rigorous, one may decompose $\phi(t)u_1$ into

$$(P_{|\tau|\le N_2}\phi(t))u_1 + (P_{|\tau|>N_2}\phi(t))u_1, \tag{5.13}$$

where the first term corresponds to frequency localization at $|\tau_1 - n_1^2| \le N_2 \ll N_2^2$, and hence the above argument can be applied. For the second term, one simply observes that

$$\|P_{|\tau|>N_2}\phi(t)\|_{L_t^\infty} \lesssim N_2^{-100}. \tag{5.14}$$

This concludes the proof.

6 Proof of Proposition 2.1: case by case study, case (*a*)

In this section, we consider case (a): $N_2(I) \lesssim N_1(II)$, $N_2(I) \ge N_3(II)$. We aim to prove for all w_1, v_2, w_3 satisfying

$$\|w_1\|_{X^{s_0,b_0}} \lesssim 1, \qquad \|w_3\|_{X^{s_0,b_0}} \lesssim 1, \qquad v_2 = \phi(t)e^{it\Delta_\gamma}\left(\sum_{n_2}\frac{g_{n_2}(\omega)}{|n_2|}e^{in_2\cdot x}\right),$$

and $\|h\|_{X^{0,1-b_0}} \lesssim 1$ that, up to an exceptional set

$$\left|\sum_{N_2\lesssim N_1, N_2\ge N_3} N_1^{s_0}\int \mathcal{N}_1(P_{N_1}w_1, P_{N_2}v_2, P_{N_3}w_3)\bar{h}\phi(t/\delta)\right| \lesssim \delta^{\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0. \tag{6.1}$$

Fix $N_{2,0}^{100} = \frac{1}{\delta}$ and recall that any loss of $\delta^{-C\epsilon_0}$ will be irrelevant in the analysis. The values of the parameters $\epsilon_0 \ll s_1 \ll s_0$ will be determined later.

By dropping a set of probability $e^{-\delta^{-cs_1}}$, we will assume the following throughout the whole section:

$$\begin{cases}|g_n(\omega)| \le N_{2,0}, & |n| \le N_{2,0},\\ |g_n(\omega)| \le |n|^{s_1}, & |n| > N_{2,0}.\end{cases} \tag{6.2}$$

And one has in particular

$$\begin{cases} \|P_{N_2}\phi(t)v_2\|_{X^{0,b_0}} \lesssim N_{2,0}^{s_1}, & N_2 \le N_{2,0}, \\ \|P_{N_2}\phi(t)v_2\|_{X^{0,b_0}} \lesssim N_2^{s_1}, & N_2 > N_{2,0}. \end{cases} \tag{6.3}$$

6.1 Standard reduction

The goal in this subsection is to reduce the estimates of this case to Lemma 6.4 and Lemma 6.5, which will be stated at the end of this subsection.

Note that in the discussion of all the cases (b)–(l), there will be a similar reduction argument. We will present the full details of the reduction in this case, and only sketch it in other cases.

We first split the summation $\sum_{N_2\lesssim N_1, N_2\ge N_3}$ into two parts $N_2 \le N_{2,0}$ and $N_2 > N_{2,0}$.

6.1.1 The low frequency part: $N_2 \le N_{2,0}$

We aim to prove

$$\sum_{N_2\le N_{2,0}, N_2\lesssim N_1, N_2\ge N_3} N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}w_1, P_{N_2}v_2, P_{N_3}w_3)\bar{h}\phi(t/\delta) \right| \lesssim \delta^{\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0. \tag{6.4}$$

Observe that, when $N_1 \gg N_2$, one can replace the h in (6.4) by $P_{N_1}h$, and when $N_1 \sim N_2$, h can be replaced by $P_{<N_1}h$.

Thus, via estimate (3.17) and (6.3), one has

- If $N_1 \sim N_2$ (in particular, $N_1 \lesssim N_{2,0}$),

$$\begin{aligned} &N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}w_1, P_{N_2}v_2, P_{N_3}w_3)\bar{h}\phi(t/\delta) \right| \\ &\lesssim \delta^{s_1/8} N_2^{2s_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} \|P_{N_2}v_2\|_{X^{0,b_0}} \|P_{N_3}w_3\|_{X^{0,b_0}} \|h\|_{X^{0,1-b_0}} \\ &\lesssim \delta^{s_1/8} N_{2,0}^{4s_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} N_1^{-s_1}. \end{aligned} \tag{6.5}$$

- If $N_1 \gg N_2$,

$$\begin{aligned} &N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}w_1, P_{N_2}v_2, P_{N_3}w_3)\bar{h}\phi(t/\delta) \right| \\ &\lesssim \delta^{s_1/8} N_2^{2s_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} \|P_{N_2}v_2\|_{X^{0,b_0}} \|w_3\|_{X^{0,b_0}} \|P_{N_1}h\|_{X^{0,1-b_0}} \\ &\lesssim \delta^{s_1/8} N_2^{3s_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} \|P_{N_1}h\|_{X^{0,1-b_0}}. \end{aligned} \tag{6.6}$$

The desired estimate will follow if one sums over the associated N_1, N_2, N_3 and apply Cauchy inequality in the sum on N_1.

Remark 6.1 We point out that the low frequency case is always the easier part in random data problems, and essentially follows from deterministic estimates usually used in the local well-posedness argument, we will not repeat this part in the rest of the article.

6.1.2 Reduction to resonant part

Now we are left with the case $N_2 > N_{2,0}$, we aim to prove

$$\sum_{N_1 \geq N_2 \geq N_3, N_2 > N_{2,0}} N_1^{s_0} \left| \int \phi(t/\delta) \bar{h} \mathcal{N}_1(P_{N_1}\phi(t/\delta)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t/\delta)w_3)\, dxdt \right| \lesssim N_{2,0}^{-\epsilon_1} \tag{6.7}$$

for some $\epsilon_1 \gg \epsilon_0$.

We will not explore the time localization $\phi(t/\delta)$ in this part. Observe that $\phi(t)h = \phi(t)\phi(t/\delta)h$, we may hence define $\tilde{h}$ as $\phi(t/\delta)h$ and use $\phi(t)\tilde{h}$ in the following estimate. Note that we still have $\|\tilde{h}\|_{X^{0,1-b_0}} \lesssim 1$. For the sake of brevity, we still denote $\tilde{h}$ as h.

Our aim is to prove for fixed N_1, N_2, N_3 satisfying $N_2 \lesssim N_1$, $N_2 \geq N_3$, $N_2 > N_{2,0}$ that

- If $N_1 \gg N_2$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1}\phi(t)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t)w_3)\, dxdt \right| \lesssim N_2^{-\epsilon_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} \|P_{N_1}h\|_{X^{0,1-b_0}}, \tag{6.8}$$

- if $N_1 \sim N_2$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1}\phi(t)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t)w_3)\, dxdt \right| \lesssim N_2^{-\epsilon_1} \|P_{N_1}w_1\|_{X^{s_0,b_0}} \|P_{<N_1}h\|_{X^{0,1-b_0}}. \tag{6.9}$$

We will focus on the proof of (6.8), and it will be easy to see that (6.9) follows similarly (almost line by line).

Observe that, since $N_1 \gg N_2 \geq N_3$, one has

$$\begin{aligned} &N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1}\phi(t)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t)w_3)\, dxdt \right| \\ &= N_1^{s_0} \left| \int \phi(t) \overline{P_{N_1}h} \mathcal{N}_1(P_{N_1}\phi(t)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t)w_3)\, dxdt \right|. \end{aligned} \tag{6.10}$$

To carry on the proof of (6.8), we introduce another parameter $M = N_2^{100s_1}$. One may split w_i, $i = 1, 3$, and v_2 as

$$\begin{aligned} \phi(t)w_i &:= P_{|\tau_i - n_i^2|<M}\phi(t)w_i + P_{|\tau_i - n_i^2|>M}\phi(t)w_i, \\ \phi(t)v_2(t) &= P_{|\tau|<M}\phi(t)v_2 + P_{|\tau|>M}\phi(t)v_2, \end{aligned} \tag{6.11}$$

and the same for h.

Remark 6.2 with such a splitting one may lose the time localization. This can be overcome by writing for example $P_{<M}\phi(t)$ as $\tilde{\phi}(t)P_{<M}\tilde{\tilde{\phi}}$, such that $\phi, \tilde{\phi}$ are Schwarz uniform in M. Or, one may further require $\phi(t) = \tilde{\psi}(t)^4$, and indeed split as $\phi(t) = \phi_1(t) + \phi_2(t)$ where $\phi_1(t) = |P_{|\tau|<M}\tilde{\psi}(t)|^4$. To make the proof clean, we leave further details to the interested reader, and allow ourselves to freely multiplying an extra time localization $\psi(t)$ in the proof.

Via (6.11), one can naturally split the left hand side of (6.8) into 2^4 parts. Each part is of the form $N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}f_1, P_{N_2}f_2, P_{N_3}f_3)\overline{P_{N_1}f_4}\right|$, where $f_i = P_{|\tau_i - n_i^2|<M}\phi(t)w_i$, or $P_{|\tau_i - n_i^2|>M}\phi(t)w_i$ for $i = 1, 3$, $f_2 = P_{|\tau|<M}\phi(t)v_2$ or $P_{|\tau|>M}\phi(t)v_2$, and $f_4 = P_{|\tau - n^2|<M}\phi(t)h$ or $P_{|\tau-n^2|>M}\phi(t)h$.

Then, applying (3.20), one has for some Schwartz function $\psi(t)$ that

$$\begin{aligned} & N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}f_1, P_{N_2}f_2, P_{N_3}f_3)\overline{P_{N_1}f_4}\right| \\ & \lesssim N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}f_1, P_{N_2}f_2, P_{N_3}f_3)\overline{P_{N_1}f_4}\psi(t)\right| \\ & \lesssim N_2^{s_1}\|P_{N_1}f_1\|_{X^{s_0,1/3}}\|P_{N_2}f_2\|_{L^\infty_{t,x}}\|P_{N_3}f_3\|_{X^{0,1/3}}\|f_4\|_{X^{0,1/3}} \\ & \quad + \|P_{N_1}f_1\|_{X^{s_0,1/3}}\|P_{N_2}f_2\|_{X^{0,b_0}}\|P_{N_2}f_3\|_{X^{0,1/3}}\|P_{N_1}f_4\|_{X^{0,1/3}}. \end{aligned} \tag{6.12}$$

(In the second line, we add a time localization $\psi(t)$, following Remark 6.2. Also recall we have $s_1 \gg \epsilon_0$.)

Unless $f_i = P_{|\tau_i - n_i^2|<M}\phi(t)w_i$, $i = 1, 3$, $f_4 = P_{<M}\phi(t)h$, and $f_2 = P_{|\tau|<M}\phi(t)v_2$, at least one of the following estimates will be true (after dropping an extra set of probability $e^{-N_2^c}$ if necessary):

- $\|P_{N_1}f_1\|_{X^{s_0,1/3}} \lesssim N_2^{-10s_1}\|P_{N_1}f_1\|_{X^{s_0,b_0}}$,
- $\|P_{N_2}f_2\|_{L^\infty_{t,x}} + \|P_{N_2}f_2\|_{X^{0,b_0}} \lesssim N_2^{-10s_1}$,
- $\|P_{N_3}f_3\|_{X^{0,1/3}} \lesssim N_2^{-10s_1}\|P_{N_3}f_3\|_{X^{0,b_0}}$,
- $\|P_{N_1}f_4\|_{X^{0,1/3}} \lesssim N_2^{-10s_1}\|P_{N_1}f_4\|_{X^{0,1-b_0}}$,

and we always have

$$\|P_{N_2}f_2\|_{X^{0,b_0}} + \|P_{N_2}f_2\|_{L^\infty_{t,x}} \lesssim N_2^{s_1}. \tag{6.13}$$

The desired estimate follows by inserting the above ones into (6.12).

Remark 6.3 The numerology in the above calculation is in fact very simple modulo lower order terms. The term $\int \mathcal{N}_1(P_{N_1}u_1, P_{N_2}u_2, P_{N_3}u_3)h\psi(t)$ can essentially be thought as $\int P_{N_1}u_1\overline{P_{N_2}u_2}P_{N_3}u_3\bar{h}\psi(t)$, and will only miss the desired estimate by at most a factor $N_2^{10s_1}$ via (3.14). On the other hand, when there is some $u_i = v_i$, which is hence already essentially localized at $|\tau - n^2| \lesssim 1$, then for all the rest of the functions h, u_j, one can gain at least $1/2 - \epsilon_0 - 1/3$ derivative. Therefore, unless all the other terms have space-time frequency localization in $|\tau - n^2| < N_2^{100s_1}$, the desired estimate will automatically follow.

Now, we are left with the case where $f_i = P_{|\tau_i - n_i^2|<M}\phi(t)w_i$, $i = 1, 3$, $f_4 = P_{|\tau-n^2|<M}\phi(t)h$, and $f_2 = P_{|\tau|<M}\phi(t)v_2$.

Let $d_1(n,t), r_2(n,t), d_3(n,t), H(n,t)$ be the space Fourier transform of $P_{|\tau_1-n_1^2|<M}\phi(t)w_1$, $P_{|\tau|<M}\phi(t)v_2$, $P_{|\tau_3-n_3^2|<M}\phi(t)w_3$, $P_{|\tau-n^2|<M}\phi(t)P_{N_1}h$. We abbreviate $d_1(n,t), r_2(n,t), d_3(n,t), H(n,t)$ as $d_1(n), r_2(n), d_3(n), H(n)$ respectively. Observe that

$$\begin{cases} N_i^{2s_0}\sum_{n\sim N_i}\|d_i(n)e^{-in_i^2t}\|_{H_t^{b_0}} \lesssim \|P_{N_1}w_i\|^2_{X^{s_0,b_0}}, \\ r_2(n_2,t) = \psi(t)\frac{g_{n_2}}{|n_2|}, \text{ for some Schwartz function } \psi, \\ \sum_{|n|\sim N_1}\|H(n)e^{-in^2t}\|_{H_t^{1-b_0}} \lesssim \|P_{N_1}h\|^2_{X^{0,1-b_0}}. \end{cases} \tag{6.14}$$

(One may observe, for example, that $\|d_i(n,t)e^{-in^2t}\|_{H_t^{b_0}} \sim \|d_i(n,t)e^{in\cdot x}\|_{X^{s_0,b_0}}$. We also point out that we have estimated $P_{|\tau|<M}\phi(t)$ just as some Schwartz function $\psi(t)$. Furthermore, one may observe that $\|d_i(n,t)e^{-in^2t}\|_{L_t^p} = \|d_i(n,t)\|_{L_t^p}$.) We will show that

$$\begin{aligned} &N_1^{s_0}\left|\int \phi(t)\bar{h}\mathcal{N}_1(P_{N_1}\phi(t)w_1, P_{N_2}\phi(t)v_2, P_{N_3}\phi(t)w_3)\,dxdt\right| \\ &= N_1^{s_0}\left|\sum_{|n_i|\sim N_i, n_1-n_2+n_3=n}\int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right| \\ &\lesssim N_2^{-\epsilon_1}\|P_{N_1}w_1\|_{X^{s_0,b_0}}\|P_{N_1}h\|_{X^{0,1-b_0}}\|P_{N_3}w_3\|_{X^{s_0,b_0}}, \text{ for some } \epsilon_1 \gg \epsilon_0. \end{aligned} \tag{6.15}$$

Observe further that $d_1(n_1,t)$, $r_2(n_2,t)$, $d_3(n_3,t)$, and $H(n,t)$ are Fourier supported in $|\tau_i - n_i^2| \lesssim M, i = 1,2,3$ and $|n-\tau^2| \lesssim M$. Thus for the integral $\int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt$ to be non-zero, one necessarily has

$$n_1^2 - n_2^2 + n_3^2 - n^2 = O(N_2^{100s_1}). \tag{6.16}$$

We thus have

$$\begin{aligned} & N_1^{s_0} \left| \sum_{|n_i|\sim N_i, n_1-n_2+n_3=n} \int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt \right| \\ & \le N_1^{s_0} \sum_{\substack{|n_i|\sim N_i, n_1-n_2+n_3-n_4=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}} \left| \int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt \right|. \end{aligned} \tag{6.17}$$

To summarize, to prove (6.8), we are left with showing the following:

Lemma 6.4 *Let $N_1 \gg N_2 \ge N_3$, then one has for some $\epsilon_1 \gg \epsilon_0$ that*

$$\begin{aligned} & N_1^{s_0} \sum_{\substack{|n_i|\sim N_i, n_1-n_2+n_3-n_4=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}} \left| \int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt \right| \\ & \lesssim N_2^{-\epsilon_1} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{N_1} h\|_{X^{0,1-b_0}} \|P_{N_3} w_3\|_{X^{s_0,b_0}}. \end{aligned} \tag{6.18}$$

We also write down the corresponding lemma that will imply (6.9).

Lemma 6.5 *Let $N_1 \sim N_2 \ge N_3$, then the same estimate (6.18) holds if one replaces the $P_{N_1}h$ by $P_{<N_1}h$.*

One can easily check that the proof of Lemma 6.4 also works for Lemma 6.5 (almost line by line).

6.2 Random data type estimate: Proof of Lemma 6.4

Recall that we always assume (6.2) and that we are in the regime $N_1 \gg N_2 \ge N_3$.

First note that for all $n_3 \sim N_3$, we have

$$\begin{aligned}\|d_3(n_3)\|_{L_t^\infty} &= \|d_3(n_3)e^{-in_3^2 t}\|_{L_t^\infty} \lesssim \|d_3(n_3)e^{-in_3^2 t}\|_{H_t^{b_0}} \sim \|d_3(n_3)e^{in_3\cdot x}\|_{X^{0,b_0}}\\ &\leq \|P_{N_3} w_3\|_{X^{0,b_0}}.\end{aligned}$$

Now, for all $|n_3| \sim N_3$ fixed, we have

$$\begin{aligned}&\sum_{\substack{|n_i|\sim N_i, i=1,2,\, n_1-n_2+n_3-n=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}} \left|\int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|\\ &\lesssim \|P_{N_3} w_3\|_{X^{0,b_0}} \sum_{\substack{|n_i|\sim N_i, i=1,2,\, n_1-n_2+n_3-n=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}} \|d_1(n_1)\|_{L_t^2}\|r_2(n_2)\|_{L_t^\infty}\|H(n)\|_{L_t^2}\\ &\lesssim \|P_{N_3} w_3\|_{X^{0,b_0}} \left(\sum_{|n|\sim N_1}\|H(n)\|_{L_t^2}^2\right)^{1/2}\\ &\quad\times\left(\sum_{|n|\sim N_1}\Big\{\sum_{\substack{|n_i|\sim N_i, i=1,2,\, n_1-n_2+n_3-n=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}}\|d_1(n_1)\|_{L_t^2}\|r_2(n_2)\|_{L_t^\infty}\Big\}^2\right)^{1/2}\\ &\lesssim \|P_{N_3} w_3\|_{X^{0,b_0}}\|P_{N_1} h\|_{X^{0,1-b_0}}\\ &\quad\times\left(N_2^{1+100s_1}\sum_{\substack{|n_i|\sim N_i, i=1,2,\, n_1-n_2+n_3-n=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}}\{\|d_1(n_1)\|_{L_t^2}\|r_2(n_2)\|_{L_t^\infty}\}^2\right)^{1/2}\\ &\lesssim \|P_{N_3} w_3\|_{X^{0,b_0}}\|P_{N_1} h\|_{X^{0,1-b_0}}\\ &\quad\times\left(N_2^{1+100s_1}\sum_{|n_i|\sim N_i, i=1,2, \langle n_2-n_1, n_2-n_3\rangle=O(N_2^{100s_1})}\|d_1(n_1)\|_{L_t^2}^2 N_2^{-2+2s_1}\right)^{1/2}.\end{aligned} \tag{6.19}$$

In the second step above, we used Cauchy inequality in n, while in the second to last step, we used Lemma 4.3. Observe as well that $\|H(n)\|_{L_t^2} \leq \|H(n)e^{-int}\|_{H_t^{1-b_0}}$, $\|r_2(n_2)\|_{L_t^\infty} \lesssim |N_2|^{s_1-1}$ up to an exceptional set, and $\|d_1(n_1)\|_{L_t^2} \leq \|d_1(n_1)e^{-in_1^2 t}\|_{L_t^2}$.

Furthermore, by the counting lemma (Lemma 4.4), one has for n_1, n_3 fixed that

$$\sharp\{|n_2| \sim N_2 : \langle n_2-n_1, n_2-n_3\rangle = O(N_2^{100s_1})\} \lesssim N_2^{100s_1}\max\left(\frac{N_2}{N_1^{1/3}}, 1\right). \tag{6.20}$$

To summarize, we derived that

$$
\begin{aligned}
&N_1^{s_0} \sum_{\substack{|n_i|\sim N_i, n_1-n_2+n_3-n_4=0,\\ n_1^2-n_2^2+n_3^2-n^2=O(N_2^{100s_1})}} \left| \int d_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\, dt \right| \\
&\lesssim N_1^{s_0} N_3^2 \|P_{N_3} w_3\|_{X^{0,b_0}} \|P_{N_1} h\|_{X^{0,1-b_0}} \\
&\quad \left(N_2^{-1+2s_1} \sum_{|n_1|\sim N_1} \|d_1(n_1)\|_{L_t^2}^2 N_2^{200s_1} \max(\frac{N_2}{N_1^{1/3}}, 1) \right)^{1/2} \\
&\lesssim N_3^2 N_2^{Cs_1} N_2^{-1/2} \left[\max\left(\frac{N_2}{N_1^{1/3}}, 1 \right) \right]^{1/2} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{N_1} h\|_{X^{0,1-b_0}} \|P_{N_3} w_3\|_{X^{s_0,b_0}}.
\end{aligned}
\tag{6.21}
$$

It is easy to see that the desired estimate will follow if there holds $N_2^{1/100} \geq N_3$.

When $N_3 \geq N_2^{\frac{1}{100}}$, we may directly go back to (6.8). Applying (3.15) and using (6.3), we have

$$
\begin{aligned}
&N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1}\phi(t) w_1,\, P_{N_2}\phi(t) v_2,\, P_{N_3}\phi(t) w_3)\, dx dt \right| \\
&\lesssim N_2^{Cs_1} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{N_2}\phi(t) v_2\|_{X^{0,b_0}} N_3^{-s_0} \|P_3 w_3\|_{X^{s_0,b_0}} \|P_{N_1} h\|_{X^{0,1-b_0}} \\
&\lesssim N_2^{Cs_1} N_3^{-s_0} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{N_1} h\|_{X^{0,1-b_0}}.
\end{aligned}
\tag{6.22}
$$

Estimate (6.8) then follows since $N_3 \geq N_2^{1/100}$ and $s_1 \ll s_0$.

7 Proof of Proposition 2.1: case by case study, case (*c*)

In this case, we have $N_1(I) \gtrsim N_2(II) \geq N_3(II)$, and we aim to prove for all v_1, w_2, w_3 satisfying

$$
v_1 = \phi(t) \sum_{n_1} \frac{g_{n_1}(\omega)}{|n_1|} e^{in_1 x}, \quad \|w_2\|_{X^{s_0,b_0}} \lesssim 1, \quad \|w_3\|_{X^{s_0,b_0}} \lesssim 1,
$$

and $\|h\|_{X_{0,1-b_0}} \lesssim 1$, that (up to an exceptional set)

$$
N_1^{s_0} \left| \int \phi(t/\delta) \bar{h} \mathcal{N}_1(P_{N_1} v_1,\, P_{N_2} w_2,\, P_{N_3} w_3)\, dx dt \right| \lesssim \delta^{\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon. \tag{7.1}
$$

7.1 Standard reduction: a (detailed) sketch

We first sketch a reduction, with an argument similar to the one in Sect. 6.1. There is indeed some difference between the reduction process in case (a) and case (c), mainly due to the difference of the form of the first term (with the highest frequency). Hence,

we will still provide a rather detailed sketch. In all the remaining cases, we only briefly sketch the reduction.

We may fix $N_{1,0}$ large satisfying $\frac{1}{\delta} = N_{1,0}^{100}$. By dropping a set of probability up to $e^{-N_{1,0}^{-cs_1}}$, we have

$$\begin{cases} |g_n(\omega)| \le N_{1,0}^{s_1}, & |n| \le N_{1,0}, \\ |g_n(\omega)| \le N_1^{s_1}, & |n| \sim N_1 \ge N_{1,0}. \end{cases} \tag{7.2}$$

By further dropping a set of probability $e^{-N_{1,0}^{-cs_1}}$ if necessary, one has

$$\begin{cases} \|P_{N_1}\phi(t)v_1\|_{X^{0,b_0}} \lesssim N_{1,0}^{s_1}, & N_1 \le N_{1,0}, \\ \|P_{N_1}\phi(t)v_1\|_{X^{0,b_0}} + \|P_{N_1}\phi(t)v_1\|_{L^\infty_{t,x}} \lesssim N_1^{s_1}, & N_1 \ge N_{1,0}. \end{cases} \tag{7.3}$$

We will assume (7.2) and (7.3) thoughout this section. Now, split into two parts $N_1 > N_{1,0}$ and $N_1 \le N_{1,0}$. For the low frequency part $N_1 \le N_{1,0}$, we may use (7.3) and apply the deterministic estimate (3.17), one thus derives the analogues of (6.5) and (6.6) below

$$\begin{aligned} & N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}v_1, P_{N_2}w_2, P_{N_3}w_3)\bar{h}\phi(t/\delta) \right| \\ & \lesssim \delta^{s_0/8} N_{1,0}^{s_0+C\epsilon_0} N_2^{s_0} \|\phi(t)v_1\|_{X^{0,b_0}} \|w_2\|_{X^{s_0,b_0}} \|w_3\|_{X^{s_0,b_0}} \|h\|_{X^{0,1-b_0}} \\ & \lesssim \delta^{s_0/8} N_{1,0}^{s_0+C\epsilon_0} N_2^{s_0} N_{1,0}^{s_1}. \end{aligned} \tag{7.4}$$

(Note that here we only need one estimate rather than two estimates as in (6.5), (6.6).)

Summing over $N_1 \le N_{1,0}$ and the associated N_2, N_3, and using the fact that $\delta^{-1} = N_{1,0}^{100}$, we derive the desired estimate

$$\sum_{N_1 \le N_{1,0}, N_1 \gtrsim N_2 \ge N_3} N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}v_1, P_{N_2}w_2, P_{N_3}w_3)\bar{h}\phi(t/\delta) \right| \lesssim \delta^{\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0. \tag{7.5}$$

For the remaining part $N_1 > N_{1,0}$, we will write $\phi(t/\delta)h$ as $\phi(t)\phi(t/\delta)h$ and note that one still has $\|\phi(t/\delta)h\|_{X^{0,1-b_0}} \lesssim 1$. For notational convenience, we will still denote $\phi(t/\delta)h$ by h, and will prove for all $N_1 \gtrsim N_2 \ge N_3$ with $N_1 > N_{1,0}$ that

$$N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1}v_1, P_{N_2}w_2, P_{N_3}w_3)\bar{h}\phi(t) \right| \lesssim N_1^{-\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0. \tag{7.6}$$

Then (7.1) will follow from summing (7.6) over N_1, N_2, N_3.

To see (7.6), we first introduce a parameter $M = N_1^{100s_0}$.

Remark 7.1 If one wants to get a rather large $s_0 < 1$, one may need to choose M more carefully. The following argument should still be fine if one chooses $M = O(N_1^{(6+)s_0})$, where 6+ denotes any number larger than 6. However, it is unclear to us whether further improvement is possible. We don't further discuss this issue here.

As in Sect. 6.1.2, we may split the functions $\phi(t)v_1 = P_{|\tau|<M}\phi(t)v_1 + P_{|\tau|>M}\phi(t)v_1$, $\phi(t)w_i = P_{|\tau_i-n_i^2|<M}\phi(t)w_i + P_{|\tau_i-n_i^2|>M}\phi(t)w_i$, $i = 2, 3$, and $\phi(t)h = P_{|\tau-n^2|<M}\phi(t)h + P_{|\tau-n^2|>M}\phi(t)h$. Applying the deterministic estimate (3.19), we reduce the proof of (7.6) to the following estimate

$$N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}\psi(t)v_1, P_{N_2}P_{|\tau_2-n_2^2|\le M}\phi(t)w_2, P_{N_3}P_{|\tau_3-n_3^2|\le M}\phi(t)w_3)\overline{P_{|\tau-n^2|\le M}\phi(t)h}\right| \lesssim N_1^{-\epsilon_1}, \tag{7.7}$$

for some $\epsilon_1 \gg \epsilon_0$. Here $\psi(t) = P_{|\tau|<M}\phi(t)$ is a Schwartz function.

Write

$$\psi(t)v_1 = \sum_{n_1} r_1(n_1,t)e^{in_1\cdot x},\ \ P_{|\tau_i-n_i^2|}\phi(t)w_i = \sum_{n_i} d_i(n_i,t)e^{in_i\cdot x},\ i = 2,3,$$
$$P_{|\tau-n^2|<M}\phi(t)h = \sum_n H(n,t)e^{in\cdot x}, \tag{7.8}$$

and abbreviate the coefficients as $r_1(n_1)$, $d_i(n_i)$ and $H(n)$ as before, one has

$$N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}\psi(t)v_1, P_{N_2}P_{|\tau_2-n_2^2|\le M}\phi(t)w_2, P_{N_3}P_{|\tau_3-n_3^2|\le M}\phi(t)w_3)\overline{P_{|\tau-n^2|\le M}\phi(t)h}\right| \le N_1^{s_0}\left|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{d_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|. \tag{7.9}$$

Observe that one has in this case the following estimates:

$$\begin{cases} N_i^{2s_0}\sum_{n_i}\|d_i(n_i)e^{-n_i^2t}\|_{H_t^{b_0}}^2 \lesssim \|P_{N_i}w_i\|_{X^{s_0,b_0}}^2 \lesssim 1,\ i = 2.3,\\ r_1(n_1,t) = \psi(t)\frac{g_{n_1}(\omega)}{|n_1|}e^{in_1\cdot x+in_1^2t},\ \text{where } \psi \text{ Schwartz},\\ \sum_{n\sim N_1}\|H(n)e^{-in^2t}\|_{H_t^{1-b_0}}^2 \lesssim \|P_{N_1}h\|_{X^{0,1-b_0}}^2. \end{cases} \tag{7.10}$$

We also point out that $\|f(t)e^{i\theta t}\|_{L_t^p} = \|f\|_{L_t^p}$. Thus, it remains to prove the following lemma:

Lemma 7.2 *Assuming (7.10), for $N_1 > N_{1,0}$, one has (up to an extra exceptional set of probability $e^{-N_1^c}$) that*

$$N_1^{s_0}\left|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{d_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right| \lesssim N_1^{-\epsilon_1},\ \textit{for some } \epsilon_1 \gg \epsilon_0. \tag{7.11}$$

7.2 Random data type estimate: Proof of Lemma 7.2

We derive three different estimates, which, combined together, will imply the desired bound.

First, one can directly go back to (7.6), and use estimate (3.15) and (7.3) to derive

$$\begin{aligned} & N_1^{s_0}\left|\int \mathcal{N}_1(P_{N_1}v_1,\, P_{N_2}w_2,\, P_{N_3}w_3)\bar{h}\phi(t)\right| \\ & \lesssim N_1^{s_0}N_1^{s_1+C\epsilon_0}\|P_{N_1}v_1\|_{X^{0,b_0}}N_2^{-s_0}\|P_{N_2}w_2\|_{X^{s_0,b_0}}N_3^{-s_0}\|P_{N_3}w_3\|_{X^{s_0,b_0}} \\ & \lesssim (N_1N_2^{-1}N_3^{-1})^{s_0}N_1^{Cs_1}. \end{aligned} \tag{7.12}$$

One can see easily that the same bound works for the left hand side of (7.11) as well. When $N_1 \sim N_2$, one can directly use (7.12) to derive the desire estimate unless $\ln N_3 \ll \ln N_1$. In particular, There is no need to consider the subcase $N_1 \sim N_2 \sim N_3$.

Next, by applying Cauchy inequality in n, one obtains

$$\begin{aligned} & N_1^{s_0}\left|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{d_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right| \\ & \lesssim N_1^{s_0}\left(\sum_n\|H(n)\|_{L_t^2}^2\right)^{1/2} \\ & \quad\left(\sum_n\left\|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(N_1^{100s_0})}} r_1(n_1)d_2(n_2)d_3(n_3)\right\|_{L_t^2}^2\right)^{1/2} \\ & \lesssim N_1^{s_0}\left(\sum_n\left\|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(N_1^{100s_0})}} r_1(n_1)d_2(n_2)d_3(n_3)\right\|_{L_t^2}^2\right)^{1/2} \end{aligned} \tag{7.13}$$

In all the summations below, we always have $|n_i| \sim N_i$, $n_1 - n_2 + n_3 = n$, $n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)$, $n_2 \neq n_1, n_3$, and we sometimes omit them for notational convenience.

One also observes that $n_1 - n_2 + n_3 = n$ and $n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)$ imply

$$\langle n_2 - n_1, n_2 - n_3\rangle = O(M), \quad \langle n_3 - n, n_3 - n_2\rangle = O(M). \tag{7.14}$$

We have the following further estimate:

$$\begin{aligned}
&\sum_n \Big\| \sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(N_1^{100s_0})}} r_1(n_1)d_2(n_2)d_3(n_3)\Big\|_{L_t^2}^2\\
&\lesssim \sum_n \Big(\sum_{n_2} \|d_2(n_2)\|_{L_t^\infty} \Big\| \sum_{n_1,n_3} r_1(n_1)d_3(n_3)\Big\|_{L_t^2}\Big)^2\\
&\lesssim \Big(\sum_{n_2}\|d_2(n_2)\|_{L_t^\infty}^2\Big)\sum_{n,n_2}\Big\| \sum_{n_1,n_3\neq n_2,\langle n_2-n_1,n_2-n_3\rangle=O(M)} \frac{g_{n_1}}{|n_1|}e^{in_1^2t}\psi(t)d_3(n_3)\Big\|_{L_t^2}^2,
\end{aligned} \tag{7.15}$$

where we used Cauchy inequality in n_2 in the last line. Since $\|d_2(n_2)\|_{L_t^\infty} \lesssim \|d_2(n_2)e^{in_2\cdot x}\|_{X^{0,b_0}}$, when $N_1 \gg N_2$, the above is further bounded by

$$\begin{aligned}
&\lesssim N_2^{-2s_0}\sum_{n,n_2}\Big\| \sum_{n_1,n_3\neq n_2,\langle n_2-n_1,n_2-n_3\rangle=O(M)} \frac{g_{n_1}}{|n_1|}e^{in_1^2t}\psi(t)d_3(n_3)\Big\|_{L_t^2}^2\\
&\lesssim N_2^{-2s_0}\sup_{n,n_2}\sharp\{n_3:\langle n_3-n_2,n_3-n\rangle=O(N_1^{100s_0})\}\\
&\qquad\sum_{\substack{n_i:n_2\neq n_1,n_3,\\ \langle n_3-n_2,n_3-n\rangle=O(N_1^{100s_0})}} \|r_1(n_1)d_3(n_3)\|_{L_t^2}^2\\
&\lesssim N_2^{-2s_0}N_1^{100s_0}\max\Big(\frac{N_3}{N_1^{1/3}},1\Big)\\
&\qquad\sum_{\substack{n_i:n_2\neq n_1,n_3,\\ \langle n_3-n_2,n_3-n\rangle=O(N_1^{100s_0})}} N_1^{-2+2s_1}\|d_3(n_3)\|_{L_t^2}^2\\
&\lesssim N_1^{Cs_0}N_1^{-2}\max\Big(\frac{N_3}{N_1^{1/3}},1\Big)\sup_{n_3}\sharp\{(n_1,n_2):\langle n_2-n_1,n_2-n_3\rangle=O(N_1^{100s_0}),\\
&\qquad n_2\neq n_1,n_3\}\sum_{n_3}\|d_3(n_3)\|_{L_t^2}^2\\
&\lesssim N_1^{Cs_0}N_1^{-2}\max\Big(\frac{N_3}{N_1^{1/3}},1\Big)N_1N_2.
\end{aligned} \tag{7.16}$$

In the above sequence of estimates, we used Hölder's inequality in the second line, and a variant of the counting Lemma 4.4 in the third line (note that we assume $N_1 \gg N_2$, thus one necessarily has $|n| \sim N_1$). In the last line we applied the counting Lemma 4.5. Moreover, note that the $\psi(t)$ in $r_1(n_1)$ gives enough decay in t, hence one has $\|d_i(n_i)\|_{L_t^\infty} = \|d_i(n_i)e^{-n_i^2t}\|_{L_t^\infty} \lesssim \|d_i(n_i)e^{-n_i^2t}\|_{H_t^{b_0}}$.

To summarize, when $N_1 \gg N_2$, one has the second estimate

$$\begin{aligned} & N_1^{s_0} \sum_{\substack{n_i \sim N_i, n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M) = O(N_1^{100 s_0})}} \left| \int r_1(n_1) \overline{d_2(n_2)} d_3(n_3) \overline{H(n)}\, dt \right| \\ & \lesssim N_1^{C s_0} \frac{N_2^{1/2}}{N_1^{1/2}} \max\left(\frac{N_3^{1/2}}{N_1^{1/6}}, 1 \right). \end{aligned} \tag{7.17}$$

One may also make use of the Frobenius norm that is more suitable when one deals with random data since it exploits better the independence of the random variables involved. The Frobenius norm together a version of the Cauchy–Schwarz inequality recalled in (C.7) will give the third estimate. We start from (7.13) again. By the same argument in (7.15), one has

$$\begin{aligned} & \sum_n \Big\| \sum_{\substack{n_i \sim N_i, n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(N_1^{100 s_0})}} r_1(n_1) d_2(n_2) d_3(n_3) \Big\|_{L_t^2}^2 \\ & \lesssim \sum_{n, n_3} \Big\| \sum_{n_1, n_3 \neq n_2, \langle n_2 - n_1, n_2 - n_3 \rangle = O(M)} \frac{g_{n_1}}{|n_1|} e^{i n_1^2 t} \psi(t) d_2(n_2) \Big\|_{L_t^2}^2, \end{aligned} \tag{7.18}$$

and by applying the recalled Cauchy–Schwarz inequality (C.7), we can further bound this expression by

$$\begin{aligned} & \lesssim N_3^2 \sup_{n_3} \sum_n \Big\| \sum_{n_2} A(n, n_2) d_2(n_2) \Big\|_{L_t^2}^2 \\ & \lesssim N_3^2 \sup_{n_3} \left(\max_n \left(\sum_{n_2} \| A(n, n_2) \|_{L_t^\infty}^2 \right) + \left(\sum_{n \neq n'} \Big\| \sum_{n_2} A(n, n_2) \overline{A(n', n_2)} \Big\|_{L_t^\infty}^2 \right)^{1/2} \right), \end{aligned} \tag{7.19}$$

where we defined

$$A(n, n_2) = A(n, n_2, t) = \begin{cases} r_1(n + n_2 - n_3), & \text{if } \langle n - n_3, n_2 - n_3 \rangle = O(N_1^{100 s_0}), \\ 0, & \text{otherwise.} \end{cases} \tag{7.20}$$

In the last line we also used $\sum_{n_2} \| d_2(n_2) \|_{L_t^2}^2 \lesssim 1$. For the sake of convenience, we also define

$$\sigma(n, n_2) = \begin{cases} \frac{g_{n + n_2 - n_3}}{N_1}, & \text{if } \langle n - n_3, n_2 - n_3 \rangle = O(N_1^{100 s_0}), \\ 0, & \text{otherwise} \end{cases} \tag{7.21}$$

for later use.

Remark 7.3 By dropping an extra set of probability $e^{-N_1^c}$, one can in fact estimate

$$\max_n \left(\sum_{n_2} \|A(n,n_2)\|_{L_t^\infty}^2 \right) + \left(\sum_{n \neq n'} \Big\| \sum_{n_2} A(n,n_2)\overline{A(n',n_2)} \Big\|_{L_t^\infty}^2 \right)^{1/2}$$

as

$$\max_n \left(\sum_{n_2} |\sigma(n,n_2)|^2 \right) + \left(\sum_{n \neq n'} \Big| \sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)} \Big|^2 \right)^{1/2}.$$

To see this, observe that $N_1 \sim |n_1| = (n_2 - n_3 + n|$, and for any fixed $t \in [0,1]$, the estimate of

$$\max_n \left(\sum_{n_2} |A(n,n_2)(t)|^2 \right) + \left(\sum_{n \neq n'} \Big| \sum_{n_2} A(n,n_2)(t)\overline{A(n',n_2)(t)} \Big|^2 \right)^{1/2}.$$

is just the same as

$$\max_n \left(\sum_{n_2} |\sigma(n,n_2)|^2 \right) + \left(\sum_{n \neq n'} \Big| \sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)} \Big|^2 \right)^{1/2}.$$

Now one can simply mimic the argument in the proof of Lemma 3.10 to go from a single t to a collection of $\{t_n\}$ in $[0,1]$ so that $|t_i - t_j| \leq N_1^{-3}$. and then go to $L_t^\infty[0,1]$. Then, finally, one can use the fact that there is a Schwartz function $\psi(t)$ multiplied inside each r_1 to go from $L_t^\infty[0,1]$ to $L_t^\infty(\mathbb{R})$. We omit the details.

In the following and throughout the rest of the article, we will estimate instead the term

$$\max_n \left(\sum_{n_2} |\sigma(n,n_2)|^2 \right) + \left(\sum_{n \neq n'} \Big| \sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)} \Big|^2 \right)^{1/2},$$

and we don't repeat the similar reduction in the rest of the article.

We fix n_3. Note that for each n, there are at most $\sim N_2 N_1^{100s_0}$ choices of n_2 so that $|n_2| \sim N_2$ and $\langle n_3 - n, n_3 - n_2 \rangle = O(N_1^{100s_0})$. Hence,

$$\max_n \left(\sum_{n_2} |\sigma(n,n_2)|^2 \right) \lesssim N_2 N_1^{Cs_0} N_1^{-2}. \tag{7.22}$$

For the non-diagonal term, we first observe that for all $n \neq n'$ fixed, up to an exceptional set of probability $e^{-N_1^{c\epsilon}}$, one can apply Lemma 3.9 to derive

$$\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right| \lesssim_\epsilon N_1^{\epsilon/2}\mathbb{E}\left(\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right|^2\right)^{1/2}. \tag{7.23}$$

This implies that

$$\begin{aligned}&\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right|^2\\ &\lesssim_\epsilon N_1^{\epsilon}\mathbb{E}\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right|^2\\ &\sim N_1^{\epsilon-4}\sharp\{n_2 : \langle n_3-n_2, n_3-n\rangle = O(N_1^{100s_0}),\\ &\quad \langle n_3-n_2, n_3-n'\rangle = O(N_1^{100s_0}),\ n_2\neq n_1, n_3\}.\end{aligned} \tag{7.24}$$

Therefore, dropping an exceptional set of probability $N_1^4 e^{-N_1^{c\epsilon}} \sim e^{-N_1^c}$, we have

$$\begin{aligned}&\sum_{n\neq n'}\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right|^2\\ &\lesssim N_1^{\epsilon-4}\sharp\{(n,n',n_2) : n\neq n',\ \langle n_3-n_2, n_3-n\rangle = O(N_1^{100s_0}),\\ &\quad \langle n_3-n_2, n_3-n'\rangle = O(N_1^{100s_0}),\ n_2\neq n_1, n_3\}.\end{aligned} \tag{7.25}$$

Counting first all the possible pairs of (n,n_2) by $N_1^{1+Cs_0}N_2$ (Lemma 4.5), and by the Wick ordered condition $n_3 \neq n_2$, which further gives at most $\sim N_1^{1+Cs_0}$ possible n', we derive

$$\left(\sum_{n\neq n'}\left|\sum_{n_2} \sigma(n,n_2)\overline{\sigma(n',n_2)}\right|^2\right)^{1/2} \lesssim N_1^{-1+Cs_0}N_2^{1/2}, \tag{7.26}$$

which obviously dominates the bound (7.22) for the diagonal term.

To summarize, we can go back to (7.18) and derive our third estimate

$$\begin{aligned}&N_1^{s_0}\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\left|\int r_1(n_1)\overline{d_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|\\ &\lesssim N_1^{Cs_0}N_1^{-1/2}N_2^{1/4}N_3.\end{aligned} \tag{7.27}$$

To complete the argument, note that the case $N_1 \sim N_2$ will follow from estimates (7.12) and (7.27). Indeed, consider two subcases. In the case $N_1 \geq N_3^{100}$, we use estimate (7.27), and when $N_1 < N_3^{100}$, we use estimate (7.12). When $N_1 \gg N_2$

(hence estimate (7.17) also holds), we also consider several subcases. In subcase $N_2N_3 \geq N_1^{11/10}$, we use estimate (7.12). In the case $N_2N_3 < N_1^{11/10}$ and $N_3 \geq N_1^{1/3}$, we use estimate (7.17). In the case $N_2N_3 < N_1^{11/10}$ and $N_3 < N_1^{1/3}$, if $N_2 \leq N_1^{9/10}$, we use estimate (7.17). Finally, if $N_2N_3 < N_1^{11/10}$, $N_3 < N_1^{1/3}$ but $N_2 > N_1^{9/10}$, there must hold $N_3 \leq N_1^{1/5}$, hence one can use estimate (7.27).

8 Proof of Proposition 2.1: Remaining cases

We present the proof of the remaining cases. Note that in each case, the desired estimate will be reduced to the resonant part similarly as in the previous two sections, and we will only briefly sketch the reduction. It is unclear whether Case (a) and Case (c) are the hardest two cases, however, all the essential arguments required to treat the rest of the cases have already appeared in the previous two sections.

We will use the following notations throughout the section. Let $\|w_i\|_{X^{s_0,b_0}} \lesssim 1$, $i = 1, 2, 3$, $v_i = \phi(t) \sum_{|n_i| \sim N_i} \frac{g_{n_i}(\omega)}{|n_i|} e^{in_i \cdot x + in_i^2 t}$, and $\|h\|_{X^{0,1-b_0}} \lesssim 1$.

Let M be a parameter that will be specified in each of the cases, $r_i(n_i, t)$ be the space Fourier transform of $(P_{|\tau| < M}\phi(t))v_i$, and $d_i(n_i, t)$ be the space Fourier transform of $P_{|\tau_i - n_i^2| < M}\phi(t)w_i$, $i = 1, 2, 3$, and $H(n, t)$ be the space Fourier transform[3] of $P_{|\tau - n^2| \leq M}\phi(t)h$. We will sometimes abbreviate $r_i(n_i, t)$, $d_i(n_i, t)$, $H(n, t)$ as $r_i(n_i)$, $d_i(n_i)$, $H(n)$ respectively.

Similarly to (6.14) and (7.10), one always has the following estimates:

$$\begin{cases} N_i^{2s_0} \sum_{n_i} \|d_i(n_i) e^{-n_i^2 t}\|^2_{H_t^{b_0}} \lesssim \|P_{N_i} w_i\|^2_{X^{s_0,b_0}} \lesssim 1, \\ r_i(n_i, t) = \psi(t) \frac{g_{n_i}(\omega)}{|n_i|} e^{in_i \cdot x + in_i^2 t}, \text{ where } \psi \text{ is a Schwartz function}, \\ \sum_{n \sim N_1} \|H(n) e^{-in^2 t}\|^2_{H_t^{1-b_0}} \lesssim \|P_{N_1} h\|^2_{X^{0,1-b_0}}. \end{cases} \tag{8.1}$$

8.1 Case (b): N_1 (II) $\geq N_3$ (I) $\geq N_2$ (II)

This part is similar to Case (a). After handling the low-frequency part using deterministic estimates and localization in time, we aim to prove for all $N_3 \geq N_{3,0}$ (where $N_{3,0}^{100} = \delta^{-1}$), one has up to an exceptional set of probability $e^{-N_3^c}$ that

- when $N_1 \sim N_3$,

$$\begin{aligned} & N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1} w_1, P_{N_2} w_2, P_{N_3} v_3) \bar{h} \phi(t) \right| \\ & \lesssim N_3^{-\epsilon_1} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{<N_1} h\|_{X^{0,1-b_0}} \\ & \sim N_1^{-\epsilon_1} \|P_{N_1} w_1\|_{X^{s_0,b_0}} \|P_{<N_1} h\|_{X^{0,1-b_0}}; \end{aligned} \tag{8.2}$$

[3] As one sees in the previous two sections, the function h here is actually $\phi(t/\delta)h$, whose $X^{0,1-b_0}$ norm is also bounded uniformly in δ.

- when $N_1 \gg N_3$,

$$N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1} w_1, P_{N_2} w_2, P_{N_3} v_3) \bar{h} \phi(t) \right| \lesssim N_3^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| P_{N_1} h \|_{X^{0,1-b_0}}. \tag{8.3}$$

Note that up to an exceptional set of probability $e^{-N_3^c}$, we can assume that

$$|g_{n_3}(\omega)| + \| P_{N_3} \phi(t) v_3 \|_{X^{0,b_0}} \leq N_3^{s_1} \tag{8.4}$$

for all $N_3 \geq N_{3,0}$, $|n_3| \sim N_3$.

It also suffices to assume

$$N_3 \geq N_2^{1000}. \tag{8.5}$$

Indeed, if $N_3 < N_2^{1000}$, from the deterministic estimate (3.15) and the bound (8.4), one obtains

$$N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1} w_1, P_{N_2} w_2, P_{N_3} v_3) \bar{f} \phi(t) \right| \lesssim N_3^{Cs_1} N_2^{-s_0} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| f \|_{X^{0,1-b_0}}, \tag{8.6}$$

where $f = P_{N_1} h$ or $P_{<N_1} h$, hence (8.2) and (8.3) follow.

In the following, we will only prove (8.2), as estimate (8.3) follows similarly (almost line by line). Note that in all the summations below we always have $|n_i| \sim N_i$, which we sometimes omit from the notation. Let $M = N_3^{100 s_1}$, similarly as in Case (a), one can reduce (8.2) to the following estimate:

$$\begin{aligned} N_1^{s_0} \left| \sum_{\substack{n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n_4^2 = O(M)}} \int d_1(n_1) \overline{d_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \\ \lesssim N_3^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| P_{<N_1} h \|_{X^{0,1-b_0}} \| P_{N_2} w_2 \|_{X^{s_0,b_0}}. \end{aligned} \tag{8.7}$$

To see this, note that one automatically has $|n| \lesssim N_1$, and recall that $\sum_{n_2} \| d_2(n_2) \|_{L^\infty}^2 \lesssim 1$. By first applying Cauchy–Schwarz in n and then in n_2, one has

$$
\begin{aligned}
&N_1^{s_0}\left|\sum_{\substack{n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}}\int d_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|\\
&\lesssim N_1^{s_0}\|P_{<N_1}h\|_{X^{0,1-b_0}}\left(\sum_n\Big\|\sum_{\substack{n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} d_1(n_1)\overline{d_2(n_2)}r_3(n_3)\Big\|_{L_t^2}^2\right)^{1/2}\\
&\lesssim N_1^{s_0}\|P_{<N_1}h\|_{X^{0,1-b_0}}\left(\sum_{n,n_2}\Big\|\sum_{\substack{n_1-n_2+n_3=n,n_2\neq n_1,n_3,\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} d_1(n_1)r_3(n_3)\Big\|_{L_t^2}^2\right)^{1/2}.
\end{aligned}
\tag{8.8}
$$

Recall also $n_1^2-n_2^2+n_3^2-n_4^2=O(M)$ together with $n_1-n_2+n_3=n$ imply that

$$
\langle n_3-n_2,n_3-n\rangle=O(M). \tag{8.9}
$$

By applying Cauchy–Schwarz in n_3 (note that the inner sum can be viewed as over n_3 only) and recalling $\|r_3(n_3)\|_{L_t^\infty}\lesssim N_3^{-1+s_1}$ (outside an exceptional set), the above can be further bounded as

$$
\begin{aligned}
&\lesssim N_1^{s_0}\|P_{<N_1}h\|_{X^{0,1-b_0}}\left(\sum_{n,n_2,n_3}\|d_1(n_1)\|_{L_t^2}^2\right)^{1/2}N_3^{-1+s_1}\\
&\quad\left(\sup_{n,n_2}\#\{n_3:\ \langle n_3-n_2,n_3-n\rangle=O(M)\}\right)^{1/2}\\
&\lesssim \|P_{<N_1}h\|_{X^{0,1-b_0}}\|P_{N_1}w_1\|_{X^{s_0,b_0}}N_3^{Cs_1}N_3^{-2/3}\\
&\quad\left(\sup_{n_1}\#\{n_2,n_3:\ \langle n_2-n_1,n_2-n_3\rangle=O(M),\ n_2\neq n_1,n_3\}\right)^{1/2},
\end{aligned}
$$

where in the second step above, we have applied a variant of Lemma 4.4 to conclude

$$
\sup_{n,n_2}\#\{n_3:\ \langle n_3-n_2,n_3-n\rangle=O(M)\}\lesssim N_3^{2/3+100s_1}.
$$

Indeed, since $N_1\sim N_3$, after dividing into $\lesssim N_3^{100s_1}$ parts, all the n_3 lie in an annulus of radius $\sim R\lesssim N_3$ with thickness $\sim O(\frac{1}{R})$. By Lemma 4.1, there are at most $\sim R^{2/3}\lesssim N_3^{2/3}$ such points.

Furthermore, we apply Lemma 4.3 to count

$$
\sup_{n_1}\#\{n_2,n_3:\ \langle n_2-n_1,n_2-n_3\rangle=O(M),\ n_2\neq n_1,n_3\}\lesssim N_2^2N_3^{1+100s_1},
$$

which implies that

$$N_1^{s_0}\left|\sum_{\substack{n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}}\int d_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|$$
$$\lesssim \|P_{<N_1}h\|_{X^{0,1-b_0}}\|P_{N_1}w_1\|_{X^{s_0,b_0}}N_3^{Cs_1}N_2N_3^{-1/6}.$$

Recall that we have reduced to the case $N_3 \geq N_2^{1000}$, hence the desired estimate is obtained.

8.2 Case (d): $N_1(I) \geq N_3(II) \geq N_2(II)$

This case is almost identical to Case (c). By the deterministic estimates, it suffices to show for all $N_1 \geq N_{1,0}$ (where $N_{1,0} = \delta^{-1}$), one has up to an exceptional set of probability $e^{-N_1^c}$ that

$$N_1^{s_0}\left|\sum_{\substack{n_i\sim N_i,n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{d_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|$$
$$\lesssim N_1^{-\epsilon_1},\ \text{for some } \epsilon_1 \gg \epsilon_0. \tag{8.10}$$

Note that by removing an exceptional set of probability $e^{-N_{1,0}^{-cs_1}}$ if necessary, we will assume in this subsection that

$$|g_{n_1}(\omega)| + \|P_{N_1}\phi(t)v_1\|_{X^{0,b_0}} + \|P_{N_1}\phi(t)v_1\|_{L^\infty_{t,x}} \lesssim N_1^{s_1},\quad \forall |n_1|\sim N_1 > N_{1,0}. \tag{8.11}$$

There still holds the same bound (7.12) as in Case (c). Moreover, one still has (7.15), as it has nothing to do with the relative sizes of N_2, N_3, and when $N_1 \gg N_3$, the bound (7.17) still holds true as well (with the choice $M = N_1^{100s_0}$). Indeed, the only step that one needs to check here is that

$$\sup_{n,n_2}\#\{n_3 :\ \langle n_3-n_2, n_3-n\rangle = O(M)\} \lesssim N_1^{100s_0}\max\left(\frac{N_3}{N_1^{1/3}},1\right),$$

which follows from the same proof of Lemma 4.4 and the assumption that $N_1 \gg N_3$.

We claim that the desired bound follows from (7.12) and (7.17). To see this, when $N_1 \sim N_3$, if one further has $N_2 > N_3^{1/9} \sim N_1^{1/9}$, one can apply (7.12). Otherwise, $N_2 \leq N_3^{1/9}$, (7.17) suffices. When $N_1 \gg N_3$, we address two difference subcases. If we are in the subcase that $N_1^{1/3} \geq N_3$, then one automatically has $N_2 \lesssim N_1^{1/3}$

hence (7.17) implies the desired estimate. In the subcase that $N_1^{1/3} < N_3$, suppose in addition one has $N_2 N_3 > N_1^{\frac{10}{9}}$, then we apply (7.12), otherwise the desired decay in N_1 follows from (7.17). The proof of Case (d) is complete.

8.3 Case (e): N_1 (II) $\geq N_2$ (I) $\geq N_3$ (I)

By a similar reduction process as in Case (a), let $N_{2,0}$ be a large parameter satisfying $N_{2,0}^{100} = \delta^{-1}$, we will focusing on proving for all $N_2 \geq N_{2,0}$ that, up to an exceptional set of probability of $e^{-N_2^c}$ and a common exceptional set independent of N_2, with probability $e^{-N_{2,0}^c}$, we have

- when $N_1 \sim N_2$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1} w_1, P_{N_2} v_2, P_{N_3} v_3) \right| \lesssim N_2^{-\epsilon_1} \sim N_1^{-\epsilon_1}, \tag{8.12}$$

- when $N_1 \gg N_2$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1} w_1, P_{N_2} v_2, P_{N_3} v_3) \right| \lesssim N_2^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| P_{N_1} h \|_{X^{0,1-b_0}}. \tag{8.13}$$

As usual, the part $N_2 \leq N_{2,0}$ will be handled by the purely deterministic estimate (3.15), and by localizing in time $\sim N_{2,0}^{-100}$.

One may assume, by dropping a set of probability $e^{-N_{2,0}^c}$, that

$$\begin{cases} |g_n| \leq N_{2,0}^{s_1}, & |n| \leq N_{2,0}, \\ |g_n| \leq N_2^{s_1}, & |n| \sim N_2 \geq N_{2,0}. \end{cases} \tag{8.14}$$

Remark 8.1 In the original paper of Bourgain [4], Case (e) is not the hardest case, however, one should be particularly careful in our irrational setting. This is because our counting lemma in the irrational case is weaker compared to the ones in [4], hence any loss of N_1^{ϵ} will be unfavorable. Since the random data argument can gain at most a (negative) power of N_2, our counting Lemma 4.5 becomes useless in Case (e).

Remark 8.2 One should also be very careful about dropping exceptional sets of small probability when the highest frequency is of type (II). For example, in our current Case (e), all large deviation type arguments require one to drop a set of probability $e^{-N_2^c}$, thus one cannot apply random data type argument for too many times. For instance, if one drops N_1^2 different sets with probability $e^{-N_2^c}$, one immediately loses control of the total probability. Moreover, in Case (e), one also needs to sum in N_1. Therefore, it is crucial that, for a fixed N_2 and for all N_1, one can apply at most N_2^C times essentially different random data type arguments. This is an issue existing even in the rational tori case. We will add some more details along the proof for the convenience of the reader.

From the remark above, one observes that the potentially most troublesome situation will be when $\ln N_1 \gg \ln N_2$. Hence, in the following we will only focus on proving (8.13), and only briefly comment on necessary changes needed for proving (8.12).

Let $M = N_2^{100 s_1}$, we may further reduce (8.13) to the following estimate:

$$N_1^{s_0} \left| \sum_{\substack{|n_i| \sim N_i,\, n_1 - n_2 + n_3 - n = 0 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)}} \int d_1(n_1) \overline{r_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \lesssim N_2^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0, b_0}} \| P_{N_1} h \|_{X^{0,1-b_0}}. \tag{8.15}$$

By applying Cauchy–Schwarz in n, we have

$$N_1^{s_0} \left| \sum_{\substack{|n_i| \sim N_i,\, n_1 - n_2 + n_3 - n = 0 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)}} \int d_1(n_1) \overline{r_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \lesssim N_1^{s_0} \| P_{N_1} h \|_{X^{0,1-b_0}} \left(\sum_n \left\| \sum_{\substack{n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n_4^2 = O(M)}} d_1(n_1) \overline{r_2(n_2)} r_3(n_3) \right\|_{L_t^2}^2 \right)^{1/2}. \tag{8.16}$$

For the sake of brevity, in the following we oftentimes omit the condition $|n_i| \sim N_i$ in the summation. Dividing $\{n_1 : |n_1| \sim N_1\}$ into finitely overlapping balls $\{J\}$ of radius $\sim N_2$, we are left with showing for each J that, up to some exceptional set of small probability $e^{-N_2^c}$,

$$\sum_{n \in J} \left\| \sum_{\substack{n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n_4^2 = O(M)}} d_1(n_1) \overline{r_2(n_2)} r_3(n_3) \right\|_{L_t^2}^2 \lesssim N_2^{-s} \| P_J w_1 \|_{X^{0,b_0}}^2, \tag{8.17}$$

where we have observed that $n_1 \in J$ implies $n \in \tilde{J}$ (a doubling of J) and we still denote $\tilde{J}$ as J for the sake of notational convenience. Moreover, we will prove the above estimate for some $s \gg s_1$. In particular, any loss of $N_2^{C s_1}$ in the estimate will be irrelevant.

Note that for each fixed N_1, there are $\sim N_1^2 / N_2^2$ such J, hence one should be careful when applying random data type argument to avoid dropping too many exceptional sets. Observe, every time one applies large deviation type argument to estimate sums of Gaussians and multiple Gausssians, one needs to drop an exceptional set of probability $e^{-N_2^c}$, and such set, a priori may depend on J. If one naively drops all such sets, a priori one may need to drop in total a set of probability $\sim \frac{N_1^2}{N_2^2} e^{-N_2^c}$, which could be enormous when $N_1 \gg N_2$. Also recall we also need to sum for all $N_1 \geq N_2$. This

problem will even arise when one studies the problem on rational tori. We will explain how to address this issue in detail in Sect. 8.3.1, and other cases will follow similarly.

Note that for the case $N_1 \sim N_2$, the decomposition into $\{J\}$ is unnecessary.

To prove (8.17), we first define

$$A(n,n_1) = A(n,n_1)(t) = \begin{cases} \sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3 \\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} r_2(n_2) r_3(n_3), & \text{if } n_1, n \in J, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\sigma(n,n_1) = \begin{cases} \sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3 \\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} \frac{g_{n_2}(\omega) g_{n_3}(\omega)}{N_2 N_3}, & \text{if } n_1, n \in J, \\ 0, & \text{otherwise.} \end{cases}$$

Then, similarly as in Remark 7.3, one has the left hand side of (8.17) bounded by

$$\lesssim \|P_J w_1\|_{X^{0,b_0}}^2 \left[\max_{n\in J} \sum_{n_1\in J} |\sigma(n,n_1)|^2 + \left(\sum_{n\neq n'} \Big| \sum_{n_1\in J} \sigma(n,n_1)\overline{\sigma(n',n_1)} \Big|^2 \right)^{1/2} \right], \tag{8.18}$$

where we have applied (C.7) and recalled that $\sum_{n_1\in J} \|d_1(n_1)\|_{L_t^2}^2 \lesssim \|P_J w_1\|_{X^{0,b_0}}^2$.

In the following, it suffices to bound the two terms in the brackets by N_2^{-s}.

The diagonal term is easier. Note that if $n_2 \neq n_3$, for $n \in J$ fixed, one has

$$\begin{aligned} &\sum_{n_1\in J} |\sigma(n,n_1)|^2 \\ &\lesssim (N_2N_3)^{-2} N_2^{Cs_1} \sup_n \#\{(n_2,n_3) : n = n_1 - n_2 + n_3,\ n_2 \neq n_1, n_3, \\ &\quad \langle n_3 - n_2, n - n_3\rangle = O(M)\} \\ &\lesssim (N_2N_3)^{-2} N_2^{Cs_1} N_3^2 N_2 = N_2^{-1+Cs_1}. \end{aligned} \tag{8.19}$$

In the first step above, we applied Lemma 3.9 to get $|\sum g_{n_2} g_{n_3}|^2 \lesssim N_2^{Cs_1} \sum 1$, by dropping an exceptional set if necessary, and in the second step, we counted n_3 naively and then n_2 using Lemma 4.3.

We are thus left with the non-diagonal term. Expanding $\sigma(n,n_1)$ and $\sigma(n',n_1)$, our goal is to show that

$$\left(\sum_{n\neq n'} \Big| \sum_{n_1\in J} \sigma(n,n_1)\overline{\sigma(n',n_1)} \Big|^2 \right)^{1/2}$$

$$= (N_2N_3)^{-2}\left(\sum_{n\neq n'}\Big|\sum_{(*)}\overline{g_{n_2}}g_{n_3}g_{n_2'}\overline{g_{n_3'}}\Big|^2\right)^{1/2} \lesssim N_2^{-s} \tag{8.20}$$

for some number $s \gg s_1$. Here we have simplified the notation by using $(*)$ to denote the set of $(n_1, n_2, n_3, n_2', n_3')$ satisfying

$$\begin{cases} n_1 \in J, \\ n = n_1 - n_2 + n_3,\ n_2 \neq n_1, n_3,\ \langle n - n_1, n - n_3\rangle = O(M), \\ n' = n_1 - n_2' + n_3',\ n_2' \neq n_1, n_3',\ \langle n' - n_1, n' - n_3'\rangle = O(M). \end{cases} \tag{8.21}$$

In the following, we will prove (8.20) case by case.

8.3.1 Case I: n_2, n_3, n_2', n_3' are distinct

Denoting the corresponding summation in $\sum_{(*)}$ by $\sum_{(*),1}$ and applying again Lemma 3.9 up to dropping a set of measure $e^{-N_2^c}$, one has

$$\Big|\sum_{(*),1}\overline{g_{n_2}}g_{n_3}g_{n_2'}\overline{g_{n_3'}}\Big|^2 \lesssim N_2^{Cs_1}\sum_{(*),1}1.$$

Hence, denoting the corresponding contribution of Case I in the left hand side of (8.20) by $(8.20)_1$, one obtains

$$(8.20)_1 \lesssim (N_2N_3)^{-2}N_2^{Cs_1}\left(\#\{n_1, n_2, n_3, n_2', n_3' :\ (**)\}\right)^{1/2}, \tag{8.22}$$

where $(**)$ denotes the conditions

$$\begin{cases} n_1 \in J, \\ n_2 \neq n_1, n_3,\ n_2' \neq n_1, n_3', \\ \langle n_2 - n_1, n_2 - n_3\rangle = O(M),\ \langle n_2' - n_1, n_2' - n_3'\rangle = O(M). \end{cases} \tag{8.23}$$

By first counting naively $n_1 \in J$, one has

$$\#\{n_1, n_2, n_3, n_2', n_3' :\ (**)\} \lesssim N_2^2\left(N_3^2\max\left(\frac{N_2}{N_1^{1/3}}, 1\right)\right)^2,$$

where we have then counted n_3 naively, and applied Lemma 4.4 (recalling that $N_1 \gg N_2$). Therefore,

$$(8.20)_1 \lesssim N_2^{Cs_1}\max(N_1^{-1/3}, N_2^{-1}) \leq N_2^{-1/3+Cs_1}.$$

Note that in the case $N_1 \sim N_2$, the same estimate remains true, as the counting Lemma 4.4 still implies the same bound. Similarly, in Case II, III, IV, V below, Lemma 4.4 always provides the same counting result.

Before we go to the next case, we explain the issue about not dropping too many exceptional sets. This needs to be taken care of since the relation

$$\begin{cases} n_1 \in J, \\ n = n_1 - n_2 + n_3,\ n_2 \neq n_1, n_3,\ \langle n - n_1, n - n_3\rangle = O(M), \\ n' = n_1 - n_2' + n_3',\ n_2' \neq n_1, n_3',\ \langle n' - n_1, n' - n_3'\rangle = O(M). \end{cases} \tag{8.24}$$

a priori depends on J. Note that we need only to worry about the case $N_1 \gg N_2, N_3$. Without loss of generality, we may assume also $n \in J$. We write $J = a_J + B_{N_2}$, $n = a_J + m$, $n' = a_J + m'$, $n_1 = a_J + m_1$, and relation (8.24) as

$$\begin{cases} m_1 \in B_{N_2}, \\ m = m_1 - n_2 + n_3,\ n_2 \neq n_1, n_3,\ \langle m - m_1, m - n_3\rangle = -\langle m - m_1, a_J\rangle + O(M), \\ m' = m_1 - n_2' + n_3',\ n_2' \neq n_1, n_3',\ \langle m' - m_1, m' - n_3'\rangle = -\langle m' - m_1, a_J\rangle + O(M). \end{cases} \tag{8.25}$$

The point is, though, there are potentially many choice of a_J, the above relation is empty unless

$$-\langle m - m_1, a_J\rangle + O(M) = O(N_2^2), \quad -\langle m' - m_1, a_J\rangle + O(M) = O(N_2^2).$$

Thus, we can always write relation (8.25) into $O(M)$ many union of the following,

$$\begin{cases} m_1 \in B_{N_2}, \\ m = m_1 - n_2 + n_3,\ n_2 \neq n_1, n_3,\ \langle m - m_1, m - n_3\rangle = a + O(1), \\ m' = m_1 - n_2' + n_3',\ n_2' \neq n_1, n_3',\ \langle m' - m_1, m' - n_3'\rangle = b + O(1), \end{cases} \tag{8.26}$$

where $a, b \in \mathbb{Z}$ and $|a|, |b| \lesssim N_2^2$. Thus, the total exceptional set one needs to drop, for all N_1 and J, will be at most $N_2^2 e^{-N_2^c}$, which is allowed.

We don't repeat this discussion of the exceptional set in the later part of the article.

8.3.2 Case II: $n_2 = n_2'$ ($n_3 \neq n_3'$)

Denote the corresponding summation in $\sum_{(*)}$ as $\sum_{(*),2}$, one has from $n_2 = n_2'$, (8.14) and Lemma 3.9 that, up to dropping an exceptional set

$$\Big| \sum_{(*),2} \overline{g_{n_2}} g_{n_3} g_{n_2'} \overline{g_{n_3'}} \Big|^2 \lesssim N_2^{Cs_1} \sum_{n_3, n_3'} (\#S(n, n', n_3, n_3'))^2, \tag{8.27}$$

where by Lemma 4.3

$$\#S(n, n', n_3, n_3') := \#\{n_1, n_2 :\ (n_1, n_2, n_3, n_2, n_3') \text{ satisfies } (*)\} \lesssim N_2^{1+Cs_1}.$$

Hence, remember the definition of $(**)$ in (8.23), one further has

$$(8.20)_2 \lesssim (N_2N_3)^{-2}N_2^{Cs_1}N_2^{1/2}\left(\#\{n_1,n_2,n_3,n_3' : (n_1,n_2,n_3,n_2,n_3') \text{ satisfies } (**)\}\right)^{1/2}. \tag{8.28}$$

By counting $n_1 \in J$ naively first, then counting (n_2,n_3) using Lemma 4.4, lastly counting n_3' via Lemma 4.3, one obtains

$$\#\{n_1,n_2,n_3,n_3' : (n_1,n_2,n_3,n_2,n_3') \text{ satisfies } (**)\} \lesssim N_2^2N_3^2\max\left(\frac{N_2}{N_1^{1/3}},1\right)N_3N_2^{Cs_1},$$

which implies that

$$(8.20)_2 \lesssim N_2^{Cs_1}N_3^{-1/2}\max(N_1^{-1/6},N_2^{-1/2}) \le N_2^{-1/6+Cs_1}. \tag{8.29}$$

8.3.3 Case III: $n_3 = n_3'$ ($n_2 \neq n_2'$)

Denoting the corresponding summation in $\sum_{(*)}$ as $\sum_{(*),3}$, similarly as in Case II, one has

$$\Big|\sum_{(*),3}\overline{g_{n_2}}g_{n_3}g_{n_2'}\overline{g_{n_3'}}\Big|^2 \lesssim N_2^{Cs_1}\sum_{n_2,n_2'}(\#S(n,n',n_2,n_2'))^2, \tag{8.30}$$

where by trivially counting n_3,

$$\#S(n,n',n_2,n_2') := \#\{n_1,n_3 : (n_1,n_2,n_3,n_2',n_3) \text{ satisfies } (*)\} \lesssim N_3^2.$$

Hence, remember the definition of $(**)$ in (8.23)

$$(8.20)_3 \lesssim (N_2N_3)^{-2}N_2^{Cs_1}N_3\left(\#\{n_1,n_2,n_2',n_3 : (n_1,n_2,n_3,n_2',n_3) \text{ satisfies } (**)\}\right)^{1/2}. \tag{8.31}$$

By trivially counting $n_1 \in J$, n_3, and applying Lemma 4.4 to n_2 and n_2', one obtains

$$\#\{n_1,n_2,n_2',n_3 : (n_1,n_2,n_3,n_2',n_3) \text{ satisfies } (**)\} \lesssim N_2^{Cs_1}N_2^2N_3^2\max\left(\frac{N_2^2}{N_1^{2/3}},1\right).$$

Therefore,

$$(8.20)_3 \lesssim N_2^{Cs_1}\max\left(N_1^{-1/3},N_2^{-1}\right) \le N_2^{-1/3+Cs_1}. \tag{8.32}$$

8.3.4 Case IV: $n_2 = n_3'$, $n_3 \neq n_2'$

Denoting the corresponding summation in $\sum_{(*)}$ as $\sum_{(*),4}$, similarly as above, one has

$$\Big|\sum_{(*),4}\overline{g_{n_2}}g_{n_3}g_{n_2'}\overline{g_{n_3'}}\Big|^2 \lesssim N_2^{Cs_1}\sum_{n_3,n_2'}(\#S(n,n',n_3,n_2'))^2, \tag{8.33}$$

where by Lemma 4.3,

$$\#S(n,n',n_3,n_2') := \#\{n_1,n_2:\ (n_1,n_2,n_3,n_2',n_2) \text{ satisfies } (*)\} \lesssim N_2^{1+Cs_1}.$$

Plugging into the above, one obtains, again remember the definition of $(**)$ in (8.23)

$$(8.20)_4 \lesssim (N_2N_3)^{-2}N_2^{Cs_1}N_2^{1/2}\left(\#\{n_1,n_2,n_3,n_2':\ (n_1,n_2,n_3,n_2',n_2) \text{ satisfies } (**)\}\right)^{1/2}. \quad (8.34)$$

Same as in Case III, by trivially counting $n_1 \in J$, n_3, and applying Lemma 4.4 to n_2 and n_2', one obtains

$$\#\{n_1,n_2,n_3,n_2':\ (n_1,n_2,n_3,n_2',n_2) \text{ satisfies } (**)\} \lesssim N_2^{Cs_1}N_2^2N_3^2\max\left(\frac{N_2^2}{N_1^{2/3}},1\right).$$

Therefore, observing that in this case one must have $N_2 \sim N_3$,

$$\begin{aligned}(8.20)_4 &\lesssim N_2^{Cs_1}N_3^{-1}N_2^{-1/2}\max\left(\frac{N_2}{N_1^{1/3}},1\right) \lesssim N_2^{-1/2+Cs_1}\max(N_1^{-1/3},N_2^{-1})\\ &\le N_2^{-5/6+Cs_1}. \end{aligned} \quad (8.35)$$

8.3.5 Case V: $n_3 = n_2'$, $n_2 \neq n_3'$

This case can be treated in the exact same way as Case IV.

8.3.6 Case VI: $n_3 = n_2'$, $n_2 = n_3'$

In this case we again have $N_2 \sim N_3$. Denoting as $\sum_{(*),6}$ the corresponding sum in $\sum_{(*)}$, one has

$$\left|\sum_{(*),6}\overline{g_{n_2}}g_{n_3}g_{n_2'}\overline{g_{n_3'}}\right|^2 \lesssim N_2^{Cs_1}\left(\#S(n,n')\right)^2, \quad (8.36)$$

where

$$S(n,n') := \{n_1,n_2,n_3:\ (n_1,n_2,n_3,n_3,n_2) \text{ satisfies } (*)\},$$

and in this case means that n_1,n_2,n_3 are distinct, $n = n_1-n_2+n_3$, $n' = n_1-n_3+n_2$, and

$$\langle n-n_1,n-n_3\rangle = O(M), \quad \langle n'-n_1,n'-n_2\rangle = O(M).$$

The above implies that $n+n' = 2n_1$, hence $\#S(n,n') \lesssim N_2^{1+Cs_1}$ by Lemma 4.3.

As a result, again remember the definition of $(**)$ in (8.23)

$$(8.20)_6 \lesssim (N_2N_3)^{-2}N_2^{Cs_1}N_2^{1/2}\left(\#\{n_1,n_2,n_3:\ (n_1,n_2,n_3,n_3,n_2) \text{ satisfies } (**)\}\right)^{1/2}. \quad (8.37)$$

From $(**)$, one has

$$\langle n_2 - n_1, n_2 - n_3\rangle = O(M), \quad \langle n_3 - n_1, n_3 - n_2\rangle = O(M),$$

hence $|n_2 - n_3|^2 = O(M)$. Trivially counting n_2, n_3, and applying Lemma 4.3 to count $n_1 \in J$, one has

$$\#\{n_1, n_2, n_3 : (n_1, n_2, n_3, n_3, n_2) \text{ satisfies } (**)\} \lesssim N_2^{Cs_1} N_2^2 N_3^2 N_2,$$

thus

$$(8.20)_6 \lesssim N_2^{Cs_1} N_3^{-1} \lesssim N_2^{-1+Cs_1}. \tag{8.38}$$

This concludes the proof of Case (e).

8.4 Case (f): $N_1(II) \geq N_3(I) \geq N_2(I)$

By the same reduction as in Case (e), let $N_{3,0}^{100} = \delta^{-1}$ and $M = N_3^{100s_1}$. It suffices to consider the high frequency part $N_3 \geq N_{3,0}$. Our goal is to show that, up to an exceptional set of probability $\sim e^{-N_3^c}$ and a common exceptional set (independent of N_3) of probability $e^{-N_{3,0}^c}$, there hold

- when $N_1 \sim N_3$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1} w_1, P_{N_2} v_2, P_{N_3} v_3) \right| \lesssim N_3^{-\epsilon_1} \sim N_1^{-\epsilon_1}, \tag{8.39}$$

- when $N_1 \gg N_3$,

$$N_1^{s_0} \left| \int \phi(t) \bar{h} \mathcal{N}_1(P_{N_1} w_1, P_{N_2} v_2, P_{N_3} v_3) \right| \lesssim N_3^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| P_{N_1} h \|_{X^{0,1-b_0}}. \tag{8.40}$$

We will also assume, by dropping a set of probability $e^{-N_{3,0}^c}$ that

$$\begin{cases} |g_n| \leq N_{3,0}^{s_1}, & |n| \leq N_{3,0}, \\ |g_n| \leq N_3^{s_1}, & |n| \sim N_3 \geq N_{3,0}. \end{cases} \tag{8.41}$$

Again, we will focus on proving (8.40). By a similar reduction as for Case (c), it suffices to show

$$\begin{aligned} N_1^{s_0} \left| \sum_{\substack{|n_i| \sim N_i,\, n_1 - n_2 + n_3 - n = 0 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)}} \int d_1(n_1) \overline{r_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \\ \lesssim N_3^{-\epsilon_1} \| P_{N_1} w_1 \|_{X^{s_0,b_0}} \| P_{N_1} h \|_{X^{0,1-b_0}}, \end{aligned} \tag{8.42}$$

which, by the same argument as in Case (e), will follow from showing for each J (of size $\sim N_3$) that, up to some exceptional set of small probability $e^{-N_3^c}$,

$$\sum_{n\in J}\Big\| \sum_{\substack{n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} d_1(n_1)\overline{r_2(n_2)}r_3(n_3)\Big\|_{L_t^2}^2 \lesssim N_3^{-s}\|P_J w_1\|_{X^{0,b_0}}^2 \tag{8.43}$$

for some constance $s \gg s_1$. We will derive two different bounds, each of which works better in different regimes of N_1, N_2, N_3.

First, by applying Hölder's inequality to the inner sum, one obtains that the left hand side of (8.43) is bounded by

$$\begin{aligned}
&\lesssim \sum_{n\in J}\sup_n \#\{n_2, n_3 : n = n_1 - n_2 + n_3,\ \langle n_3 - n_2, n - n_3\rangle = O(M)\}\\
&\quad \sum_{n_1,n_2,n_3} \|d_1(n_1)\overline{r_2(n_2)}r_3(n_3)\|_{L_t^2}^2\\
&\lesssim N_3^{Cs_1}N_2^2 \max\left(\frac{N_3}{N_1^{1/3}}, 1\right) \sum_{n,n_1,n_2,n_3} N_2^{-2}N_3^{-2}\|d_1(n_1)\|_{L_t^2}^2 \qquad (8.44)\\
&\lesssim N_3^{Cs_1}\max\left(\frac{N_3}{N_1^{1/3}}, 1\right) N_3^{-2}\|P_J w_1\|_{X^{0,b_0}}^2 \sup_{n_1}\#\{n_2, n_3 : n_2 \neq n_1, n_3,\\
&\quad \langle n_2 - n_1, n_2 - n_3\rangle = O(M)\}.
\end{aligned}$$

In the second line above, we applied (8.41) and Lemma 4.4 (the estimate holds true in both cases $N_1 \gg N_3$ and $N_1 \sim N_3$). One can then trivially count n_2 and apply Lemma 4.3 to count n_3 to further bound the above by

$$\begin{aligned}
&\lesssim N_3^{Cs_1}\max\left(\frac{N_3}{N_1^{1/3}}, 1\right) N_3^{-2}\|P_J w_1\|_{X^{0,b_0}}^2 N_2^2 N_3\\
&\leq N_3^{Cs_1}N_2^2\max(N_1^{-1/3}, N_3^{-1})\|P_J w_1\|_{X^{0,b_0}}^2. \qquad (8.45)
\end{aligned}$$

We now turn to a different estimate of the left hand side of (8.43). In fact, we claim that by the same argument as in Case (e), one also obtains in our current case that

$$\sum_{n\in J}\Big\| \sum_{\substack{n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} d_1(n_1)\overline{r_2(n_2)}r_3(n_3)\Big\|_{L_t^2}^2 \lesssim N_3^{Cs_1}N_2^{-1/6}\|P_J w_1\|_{X^{0,b_0}}^2. \tag{8.46}$$

We first explain how to complete the proof of (8.43) using the two estimates (8.45) and (8.46) above. In the case $N_2 > N_1^{\frac{1}{6}-\frac{1}{100}}$, the desired estimate follows from (8.46). Now suppose $N_2 \leq N_1^{\frac{1}{6}-\frac{1}{100}}$. If $N_3 \leq N_1^{1/3}$, the bound in (8.45) becomes $N_3^{Cs_1}N_2^2N_3^{-1}$. Hence, in the subcase when $N_2 \leq N_3^{\frac{1}{2}-\frac{1}{100}}$, this is good enough. If $N_2 > N_3^{\frac{1}{2}-\frac{1}{100}}$,

one can again apply (8.46) to obtain the desired bound. It is left to check the case $N_3 > N_1^{1/3}$, where (8.45) becomes $N_3^{Cs_1} N_2^2 N_1^{-1/3}$. Observe that this does imply the desired result, since we are already in the regime $N_2 \leq N_1^{\frac{1}{6}-\frac{1}{100}}$.

It thus suffices to verify (8.46). Note that most of the estimates in Case (e) still hold, as they do not depend on the relative sizes of N_2, N_3. More precisely, the estimate (8.19) for the diagonal term, and the final bounds obtained in Case VI, V, VI for the non-diagonal term still hold true (with an extra factor $N_3^{Cs_1}$). We are now left to examine Case I, II, III.

8.4.1 Case I: n_2, n_3, n_2', n_3' are distinct

We start with estimate (8.22), which still holds true in Case (f). Note that

$$\#\{n_1, n_2, n_3, n_2', n_3' : (**)\} = \#\{n, n_2, n_3, n_2', n_3' : (***)\},$$

where $(**)$ is in Case (e), given in (8.23), and $(***)$ denotes the conditions

$$\begin{cases} n \in J, \\ n_3 \neq n, n_2, \ n_3' \neq n, n_2', \\ \langle n_3 - n_2, n - n_3 = O(M), \quad \langle n_3' - n_2', n - n_3' \rangle = O(M). \end{cases} \tag{8.47}$$

We first count $n \in J$ naively (recalling that $|J| \sim N_3$), we then count n_2 naively and n_3 using Lemma 4.4, and repeat for (n_2', n_3'). This leads to

$$\#\{n, n_2, n_3, n_2', n_3' : (***)\} \lesssim N_3^{Cs_1} N_3^2 \left(N_2^2 \max\left(\frac{N_3}{N_1^{1/3}}, 1 \right) \right)^2.$$

Note that this bound holds true in both cases $N_1 \gg N_3$ and $N_1 \sim N_3$. Hence, one obtains

$$\begin{aligned} (8.20)_1 &\lesssim (N_2 N_3)^{-2} N_3^{Cs_1} \left(\#\{n_1, n_2, n_3, n_2', n_3' : (**)\} \right)^{1/2} \\ &\lesssim (N_2 N_3)^{-2} N_3^{Cs_1} N_2^2 N_3 \max\left(\frac{N_3}{N_1^{1/3}}, 1 \right) \\ &\lesssim N_3^{Cs_1} \max(N_1^{-1/3}, N_3^{-1}) \leq N_3^{-1/3+Cs_1} \lesssim N_3^{Cs_1} N_2^{-1/3}. \end{aligned} \tag{8.48}$$

8.4.2 Case II: $n_2 = n_2'$ ($n_3 \neq n_3'$)

In our current case, after dropping an exceptional set, one still has estimate (8.22), where

$$\begin{aligned} \#S(n, n', n_3, n_3') &:= \#\{n_1, n_2 : (n_1, n_2, n_3, n_2, n_3') \text{ satisfies } (*)\} \\ &\lesssim \#\{n_2 : \langle n_3 - n_2, n - n_3 \rangle = O(M)\} \lesssim N_3^{Cs_1} N_2. \end{aligned}$$

Hence, remembering the definition of $(**)$ in (8.23), one has

$$(8.20)_2 \lesssim (N_2N_3)^{-2}N_3^{Cs_1}N_2^{1/2}\left(\#\{n_1,n_2,n_3,n_3' : (n_1,n_2,n_3,n_2,n_3') \text{ satisfies } (**)\}\right)^{1/2}. \quad (8.49)$$

By counting n_2 first naively, then n_3 naively, then n_1 by Lemma 4.3, and lastly n_3' by Lemma 4.3 as well, one obtains

$$\#\{n_1,n_2,n_3,n_3' : (n_1,n_2,n_3,n_2,n_3') \text{ satisfies } (**)\} \lesssim N_3^{Cs_1}N_2^2N_3^2N_3N_3,$$

which implies

$$(8.20)_2 \lesssim N_3^{Cs_1}N_2^{-1/2}. \quad (8.50)$$

8.4.3 Case III: $n_3 = n_3'$ ($n_2 \neq n_2'$)

Note that after dropping an exceptional set, estimate (8.30) still holds true in Case (f). But this time, we count $\#S(n,n',n_2,n_2')$ more carefully. It suffices to count n_3 satisfying $\langle n_3 - n_2, n - n_3\rangle = O(M)$. By Lemma 4.4, one has

$$\#S(n,n',n_2,n_2') \lesssim \begin{cases} N_3^{Cs_1}N_3^{2/3}, & \text{if } N_1 \sim N_3, \\ N_3^{Cs_1}\max\left(\frac{N_3}{N_1^{1/3}},1\right), & \text{if } N_1 \gg N_3. \end{cases}$$

In the second estimate above, we used the fact that $|n| \sim N_1$ when $N_1 \gg N_3$. Same as before, where remembering the definition of $(**)$ in (8.23),

$$(8.20)_3 \lesssim (N_2N_3)^{-2}N_3^{Cs_1}\left(\#S(n,n',n_2,n_2')\right)^{1/2} \left(\#\{n_1,n_2,n_2',n_3 : (n_1,n_2,n_3,n_2',n_3) \text{ satisfies } (**)\}\right)^{1/2}. \quad (8.51)$$

By counting n_1, n_2, n_2' trivially, and then n_3 using Lemma 4.3, one has

$$\#\{n_1,n_2,n_2',n_3 : (n_1,n_2,n_3,n_2',n_3) \text{ satisfies } (**)\} \lesssim N_3^{Cs_1}N_3^2N_2^4N_3.$$

Combining the above bounds together, one obtains

$$
\begin{aligned}
(8.20)_3 &\lesssim (N_2N_3)^{-2}N_3^{Cs_1}\left(\#S(n,n',n_2,n_2')\right)^{1/2}N_3^{3/2}N_2^2\\
&\lesssim N_3^{Cs_1}N_3^{-1/2}\cdot\begin{cases} N_3^{1/3}, & \text{if } N_1\sim N_3,\\ \max\left(\frac{N_3^{1/2}}{N_1^{1/6}},1\right), & \text{if } N_1\gg N_3\end{cases}\\
&\lesssim N_3^{Cs_1}N_3^{-1/6}\le N_3^{Cs_1}N_2^{-1/6}.
\end{aligned}
\tag{8.52}
$$

The proof of Case (f) is thus complete.

8.5 Case (g): $N_1(I)\ge N_2(I)\ge N_3(II)$

By deterministic estimates, it suffices to show for all $N_1\ge N_{1,0}$ (where $N_{1,0}=\delta^{-1}$), one has up to an exceptional set of probability $e^{-N_1^c}$ that

$$
\begin{aligned}
&N_1^{s_0}\left|\sum_{\substack{n_i\sim N_i, n_1-n_2+n_3=n,\ n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|\\
&\lesssim N_1^{-\epsilon_1},\ \text{for some } \epsilon_1\gg\epsilon_0,
\end{aligned}
\tag{8.53}
$$

where $M=O(N_1^{100s_0})$.

By dropping an exceptional set of probability $e^{-N_{1,0}^c}$, one may assume that

$$
|g_{n(\omega)}|\lesssim N_1^{s_1},\quad \forall |n|\sim N_1\ge N_{1,0}. \tag{8.54}
$$

Our goal is to show that, up to an exceptional set of probability $e^{-N_1^c}$, there holds

$$
N_1^{s_0}\left|\int\phi(t)\bar{h}\mathcal{N}_1(P_{N_1}v_1,P_{N_2}v_2,P_{N_3}w_3)\right|\lesssim N_1^{-\epsilon_1}. \tag{8.55}
$$

Apparently, (8.55) can be reduced to the following estimate:

$$
N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\ n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|\lesssim N_1^{-\epsilon_1}. \tag{8.56}
$$

We first derive an estimate that will handle the regime $N_3 \le N_1^{1-\frac{1}{100}}$. To see this, applying Cauchy–Schwarz in n, one obtains

$$\begin{aligned}
&N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\, n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\overline{H(n)}\,dt\right|\\
&\lesssim N_1^{s_0}\left(\sum_n \|H(n)\|_{L_t^2}^2\right)^{1/2}\left(\sum_n\Big\|\sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\Big\|_{L_t^2}^2\right)^{1/2}\\
&\lesssim N_1^{s_0}\left(\sum_n\Big\|\sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\Big\|_{L_t^2}^2\right)^{1/2}.
\end{aligned}\tag{8.57}$$

Then, applying Cauchy–Schwarz again to the inner sum above, one has, after dropping an exceptional set of probability $e^{-N_1^c}$,

$$\begin{aligned}
&\lesssim N_1^{s_0}N_1^{-1}N_2^{-1}\left(\sum_n\Big|\sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}}\|d_3(n_3)\|_{L_t^2}\Big|^2\right)^{1/2}\\
&\lesssim N_1^{s_0}N_1^{-1}N_2^{-1}\left(\sum_n\sum_{n_1,n_2,n_3}\|d_3(n_3)\|_{L_t^2}^2\right)^{1/2}\\
&\quad\left(\sup_n \#\{n_2,n_3:\ \langle n_3-n_2, n-n_3\rangle = O(M)\}\right)^{1/2}\\
&\lesssim N_1^{Cs_0}N_1^{-1}N_2^{-1}N_2^{1/2}N_3^{1/2}\\
&\quad\left(\sup_{n_3}\#\{n_1,n_2:\ n_2\neq n_1,n_3,\ \langle n_3-n_1,n_3-n_2\rangle = O(M)\}\right)^{1/2},
\end{aligned}\tag{8.58}$$

where in the last step above, we have applied Lemma 4.5. Another application of Lemma 4.5 implies that

$$\lesssim N_1^{Cs_0}N_1^{-1}N_2^{-1}N_2^{1/2}N_3^{1/2}N_1^{1/2}N_2^{1/2}\lesssim N_1^{Cs_0}N_1^{-1/2}N_3^{1/2}.\tag{8.59}$$

Hence, the desired estimate follows if $N_3 \le N_1^{1-\frac{1}{100}}$.

The other case $N_3 > N_1^{1-\frac{1}{100}}$ in fact follows directly from the estimates in Case (e). Note that the relative sizes of N_1, N_2, N_3 in these two cases are the same, so all the

counting in Case (e) remain true here. We briefly sketch the argument here. It suffices to show, up to an exceptional set of small probability $e^{-N_1^c}$, that

$$\sum_n \Big\| \sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3 \\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} r_1(n_1)\overline{r_2(n_2)}d_3(n_3)\Big\|_{L_t^2}^2 \lesssim N_1^{-\epsilon_1}. \tag{8.60}$$

We would like to apply again version of the Cauchy–Schwarz inequality in (C.7), but this time with

$$\sigma(n,n_3) = \sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3 \\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}} \frac{g_{n_1}(\omega)g_{n_2}(\omega)}{N_1N_2}.$$

For the diagonal term, the exact same counting as in (8.19) gives, for any fixed n,

$$\sum_{n_3}|\sigma(n,n_3)|^2 \lesssim (N_1N_2)^{-2}N_1^{Cs_0}N_3^2N_2 \lesssim N_1^{Cs_0}N_2^{-1}, \tag{8.61}$$

which is good enough since $N_2 \geq N_3 > N_1^{1-\frac{1}{100}}$.

The six cases for the non-diagonal term work similarly. In particular, Case I (all n_1, n_1', n_2, n_2' distinct), Case III ($n_1 = n_1', n_2 \neq n_2'$), and Case VI ($n_1 = n_2', n_2 = n_1'$) can be carried out in the exact same way.

In Case II ($n_2 = n_2', n_1 \neq n_1'$), one has up to an exceptional set that

$$\Big|\sum_{(*),2}\overline{g_{n_1}}g_{n_2}g_{n_1'}\overline{g_{n_2'}}\Big|^2 \lesssim N_1^{Cs_0}\sum_{n_1,n_1'}(\#S(n,n',n_1,n_1'))^2 \tag{8.62}$$

Here, by Lemma 4.3,

$$\#S(n,n',n_1,n_1') := \#\{n_2,n_3 : n_1,n_2,n_3,n_1',n_2 \text{ satisfies } (*')\} \lesssim N_1^{Cs_0}N_3,$$

with $(*')$ denotes conditions

$$\begin{cases} n_1 \in J, \\ n = n_1 - n_2 + n_3,\ n_2 \neq n_1, n_3,\ \langle n-n_1, n-n_3\rangle = O(M), \\ n' = n_1' - n_2' + n_3,\ n_2' \neq n_1', n_3,\ \langle n'-n_1', n'-n_3\rangle = O(M). \end{cases} \tag{8.63}$$

The rest of the argument proceeds in the exact same way as Case II in Case (e).

Since Case IV and V can be dealt with in the same way, we only briefly discuss Case IV ($n_2 = n_1', n_1 \neq n_2'$) here. By dropping an exceptional set, one has the following analogue of(8.33):

$$\Big|\sum_{(*),4}\overline{g_{n_1}}g_{n_2}g_{n_1'}\overline{g_{n_2'}}\Big|^2\lesssim N_1^{Cs_0}\sum_{n_1,n_2'}(\#S(n,n',n_1,n_2'))^2, \tag{8.64}$$

where by Lemma 4.3,

$$\#S(n,n',n_1,n_2'):=\#\{n_2,n_3:\ (n_1,n_2,n_3,n_2,n_2')\text{ satisfies }(*')\}\lesssim N_1^{Cs_0}N_3.$$

The rest of the argument again proceeds in the same way as in Case IV of Case (e), which is left to the reader. The proof of Case (g) is complete.

8.6 Case (h): $N_1(I)\geq N_3(II)\geq N_2(I)$

This case can be estimated in the same way as Case (g), where again if $N_3>N_1^{1-\frac{1}{100}}$, the bounds in Case (f) apply. Note that, compared to Case (f), one can think about $N_1(I)$, in Case (h), as $N_1(II)$, except one suffers a loss $N_1^{2s_0}$. The computation in Case (f) gives a gain of $N_3^{-s^*}$, where s^* can be computed explicitly. Thus, when s_0 is small enough and $N_3\geq N_1^{\frac{1}{100}}$, the extra loss of $N_1^{2s_0}$ can be neglected. We omit the details.

8.7 Case (i), (j): $N_1(I)\geq N_2(II)\geq N_3(I)$ or $N_1(I)\geq N_3(I)\geq N_2(II)$

As before, we focus on showing for all $N_1\geq N_{1,0}$ (where $N_{1,0}=\delta^{-1}$), up to an exceptional set of probability $e^{-N_1^c}$, there holds

$$\begin{aligned}&N_1^{s_0}\Bigg|\sum_{\substack{n_i\sim N_i,n_1-n_2+n_3=n,\,n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n^2=O(M)=O(N_1^{100s_0})}}\int r_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\Bigg|\\ &\lesssim N_1^{-\epsilon_1},\ \text{for some }\epsilon_1\gg\epsilon_0,\end{aligned} \tag{8.65}$$

where $M=O(N_1^{100s_0})$.

We may assume dropping an exceptional set of probability $e^{-N_{1,0}^c}$ that

$$|g_{n(\omega)}|\lesssim N_1^{s_1},\quad \forall|n|\sim N_1\geq N_{1,0}. \tag{8.66}$$

We aim to show that, up to an exceptional set of probability $e^{-N_1^c}$, there holds

$$N_1^{s_0}\Big|\int\phi(t)\bar{h}\mathcal{N}_1(P_{N_1}v_1,P_{N_2}w_2,P_{N_3}v_3)\Big|\lesssim N_1^{-\epsilon_1}. \tag{8.67}$$

Similarly as before, (8.67) will follow from the following estimate:

$$N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\, n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|\lesssim N_1^{-\epsilon_1}. \tag{8.68}$$

We will first introduce an estimate that allows one to reduce to the regime $N_2 > N_1^{1-\frac{1}{100}}$. To see this, following the same Cauchy–Schwarz argument as in (8.57), (8.58), one obtains, after dropping a set of probability $e^{-N_1^c}$,

$$\begin{aligned}
&N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\, n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|\\
&\lesssim N_1^{Cs_0}N_1^{-1}N_3^{-1}N_2^{1/2}N_3^{1/2}\\
&\left(\sup_{n_2}\#\{n_1,n_3:\, n_2\neq n_1,n_3,\ \langle n_3-n_1,n_3-n_2\rangle=O(M)\}\right)^{1/2}\\
&\lesssim N_1^{Cs_0}N_1^{-1}N_3^{-1}N_2^{1/2}N_3^{1/2}N_1^{1/2}N_3^{1/2}\lesssim N_1^{Cs_0}N_1^{-1/2}N_2^{1/2},
\end{aligned} \tag{8.69}$$

where in the last two steps we have applied Lemma 4.5. Hence, if $N_2\le N_1^{1-\frac{1}{100}}$, the desired estimate follows.

Next, we may apply the same estimate as in Case (c) to deal with the case $N_3\le N_1^{1/5}$. Indeed, introducing an extra factor of $N_3^{s_0}$ (since the third input function in the current case is random) to (7.27), one has

$$N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\, n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|\lesssim N_1^{Cs_0}N_1^{-1/2}N_2^{1/4}N_3, \tag{8.70}$$

which is enough to handle the case $N_3\le N_1^{1/5}$.

To summarize, we have reduced the desired estimate to the regime $N_2 > N_1^{1-\frac{1}{100}}$ and $N_3 > N_1^{1/5}$. The rest of the argument is essentially repeating that of Case (e). Define

$$\sigma(n,n_2)=\sum_{\substack{n_1-n_2+n_3=n,\, n_2\neq n_1,n_3\\ n_1^2-n_2^2+n_3^2-n_4^2=O(M)}}\frac{g_{n_1}(\omega)g_{n_3}(\omega)}{N_1N_3}.$$

Then, one has

$$N_1^{s_0}\left|\sum_{\substack{|n_i|\sim N_i,\, n_1-n_2+n_3-n=0\\ n_1^2-n_2^2+n_3^2-n^2=O(M)}}\int r_1(n_1)\overline{d_2(n_2)}r_3(n_3)\overline{H(n)}\,dt\right|$$
$$\lesssim N_1^{Cs_0}\left[\max_n\sum_{n_2}|\sigma(n,n_2)|^2+\left(\sum_{n\neq n'}\Big|\sum_{n_2}\sigma(n,n_2)\overline{\sigma(n',n_2)}\Big|^2\right)^{1/2}\right]^{1/2}. \tag{8.71}$$

Again, the diagonal term is easier to deal with. After dropping a set of small probability and applying Lemma 4.5, one has

$$\max_n\sum_{n_2}|\sigma(n,n_2)|^2\lesssim N_1^{Cs_0}N_1^{-2}N_3^{-2}N_2N_3\lesssim N_1^{-1+Cs_0}. \tag{8.72}$$

For the non-diagonal term, rewrite

$$\left(\sum_{n\neq n'}\Big|\sum_{n_2}\sigma(n,n_2)\overline{\sigma(n',n_2)}\Big|^2\right)^{1/2}=(N_1N_3)^{-2}\left(\sum_{n\neq n'}\Big|\sum_{(*)''}g_{n_1}g_{n_3}\overline{g_{n_1'}g_{n_3'}}\Big|^2\right)^{1/2}, \tag{8.73}$$

where $(*'')$ denotes the set of (n_1,n_2,n_3,n_1',n_3') satisfying

$$\begin{cases} n=n_1-n_2+n_3,\ n_2\neq n_1,n_3,\ \langle n-n_1,n-n_3\rangle=O(M),\\ n'=n_1'-n_2+n_3',\ n_2\neq n_1',n_3',\ \langle n'-n_1',n'-n_3'\rangle=O(M). \end{cases} \tag{8.74}$$

We discuss three subcases in the following, and omit the symmetric ones. Note that they proceed very similarly as the corresponding cases in Case (e).

8.7.1 Case I: n_1, n_3, n_1', n_3' are distinct

By dropping an exceptional set, one obtains

$$(8.73)_1\lesssim N_1^{Cs_0}(N_1N_3)^{-2}(\#S)^{1/2}, \tag{8.75}$$

where S denotes the set (n_1,n_2,n_3,n_1',n_3') satisfying

$$\begin{cases} n_2\neq n_1,n_3,n_1',n_3',\\ \langle n_2-n_1,n_2-n_3\rangle=O(M),\quad \langle n_2-n_1',n_2-n_3'\rangle=O(M). \end{cases} \tag{8.76}$$

Counting n_2 trivially first, then applying Lemma 4.5 twice, one has

$$\#S\lesssim N_1^{Cs_0}N_2^2(N_1N_3)^2,$$

which implies

$$(8.73)_1 \lesssim N_1^{Cs_0} N_1^{-1} N_2 N_3^{-1} \lesssim N_1^{-1/5+Cs_0}. \tag{8.77}$$

8.7.2 Case II: $n_1 = n_1'$ ($n_3 \neq n_3'$)

Similarly as in Case II of Case (e), one obtains

$$(8.73)_2 \lesssim N_1^{Cs_0} (N_1 N_3)^{-2} \left(\#S(n, n', n_3, n_3')\right)^{1/2} (\#S)^{1/2}, \tag{8.78}$$

where remembering the definition of $(*'')$ in (8.74)

$$\#S(n, n', n_3, n_3') = \#\{n_1, n_2 : (n_1, n_2, n_3, n_1, n_3') \text{ satisfies } (*'')\} \lesssim N_1^{Cs_0} N_1$$

by Lemma 4.3, and S denotes the set of (n_1, n_2, n_3, n_3') satisfying (8.76). Hence, by counting n_2 trivially, then counting n_1, n_3 via Lemma 4.5, and finally counting n_3' according to Lemma 4.3, one obtains

$$\#S \lesssim N_1^{Cs_0} N_2 N_1 N_3 N_3.$$

Combining together, one has

$$(8.73)_2 \lesssim N_1^{Cs_0} (N_1 N_3)^{-2} N_1^{1/2} N_1^{1/2} N_2^{1/2} N_3 \lesssim N_1^{Cs_0} N_1^{-1} N_2^{1/2} N_3^{-1} \lesssim N_1^{-1/2+Cs_0}. \tag{8.79}$$

8.7.3 Case III: $n_1 = n_3'$ ($n_3 \neq n_1'$)

Following the same calculation as in Case III of Case (e), one has, up to a small exceptional set,

$$(8.73)_2 \lesssim N_1^{Cs_0} (N_1 N_3)^{-2} \left(\#S(n, n', n_3, n_1')\right)^{1/2} (\#S)^{1/2}, \tag{8.80}$$

where

$$\#S(n, n', n_3, n_1') = \#\{n_1, n_2 : (n_1, n_2, n_3, n_1', n_1) \text{ satisfies } (*'')\} \lesssim N_1^{Cs_0} N_1$$

according to Lemma 4.3. In the above, S consists of (n_1, n_2, n_3, n_1') so that $(n_1, n_2, n_3, n_1', n_1)$ satisfies (8.76). One thus has

$$\#S \lesssim N_3^2 (N_1^{Cs_0+1} N_2) N_1$$

by Lemma 4.5 and 4.3 similarly as before. Therefore,

$$(8.73)_2 \lesssim N_1^{Cs_0} (N_1 N_3)^{-2} N_1^{1/2} N_1 N_2^{1/2} N_3 \lesssim N_1^{Cs_0} N_1^{-1/2} N_2^{1/2} N_3^{-1} \lesssim N_1^{-1/5+Cs_0}. \tag{8.81}$$

This concludes the proof of Case (i) and (j).

8.8 Case (k), (l): $N_1(I) \geq N_2(I) \geq N_3(I)$ or $N_1(I) \geq N_3(I) \geq N_2(I)$

Considering only the case $N_1 \geq N_{1,0}$, where $N_{1,0}^{100} = \delta^{-1}$. Our goal is to show that

$$N_1^{s_0} \left| \int \mathcal{N}_1(P_{N_1} v_1, P_{N_2} v_2, P_{N_3} v_3) \bar{h}\phi(t) \right| \lesssim N_1^{-\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0 \tag{8.82}$$

up to an exceptional set of probability $e^{-N_1^c}$.

Let $M = O(N_1^{100 s_0})$, by a similar reduction argument as in Case (c), it suffices to prove that

$$\begin{aligned} & N_1^{s_0} \left| \sum_{\substack{n_i \sim N_i, n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)}} \int r_1(n_1) \overline{r_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \\ & \lesssim N_1^{-\epsilon_1}, \text{ for some } \epsilon_1 \gg \epsilon_0. \end{aligned} \tag{8.83}$$

Following the same argument in (8.57), one has

$$\begin{aligned} & N_1^{s_0} \left| \sum_{\substack{n_i \sim N_i, n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n^2 = O(M)}} \int r_1(n_1) \overline{r_2(n_2)} r_3(n_3) \overline{H(n)}\, dt \right| \\ & \lesssim N_1^{s_0} \left(\sum_n \Big\| \sum_{\substack{n_1 - n_2 + n_3 = n,\, n_2 \neq n_1, n_3 \\ n_1^2 - n_2^2 + n_3^2 - n_4^2 = O(M)}} r_1(n_1) \overline{r_2(n_2)} r_3(n_3) \Big\|_{L_t^2}^2 \right)^{1/2}. \end{aligned} \tag{8.84}$$

Suppose n_1, n_2, n_3 are all distinct, then by dropping a set of probability $e^{-N_1^c}$ and recalling the presence of Schwartz function $\psi(t)$ in each $r_i(n_i, t)$, one can bound the above by

$$\begin{aligned} & \lesssim N_1^{Cs_0} (N_1 N_2 N_3)^{-1} \left(\#\{n_1 \neq n_2 \neq n_3 : \langle n_2 - n_1, n_2 - n_3 \rangle = O(M)\} \right)^{1/2} \\ & \lesssim N_1^{Cs_0} (N_1 N_2 N_3)^{-1} (N_2^2 N_3^2 N_1)^{1/2} \lesssim N_1^{-1/2 + Cs_0}, \end{aligned} \tag{8.85}$$

where in the second line above we trivially counted n_2, n_3 and applied Lemma 4.3 to count n_1.

Now, suppose $n_1 = n_3$. For fixed n, one thus has $n = 2n_3 - n_2$ and $|n_3 - n_2|^2 = O(M)$, hence $|n_3 - n|^2 = O(M)$. By losing a factor of $N_1^{Cs_0}$, we can assume that $|n_3 - n|^2 = \mu + O(1)$, where $\mu \lesssim N_1^{Cs_0}$, in other words, n_3 lies in an annulus of radius $\sim R$ and thickness $\sim O(\frac{1}{R})$, with $R \lesssim N_1^{Cs_0}$. Applying Lemma 4.1, one sees that the total number of such n_3 is at most $\lesssim R^{2/3} \lesssim N_1^{Cs_0}$.

Therefore, by Cauchy–Schwarz, outside an exceptional set of probability $e^{-N_1^c}$, one has

$$\begin{aligned}(8.84) &\lesssim N_1^{Cs_0}\,(\#\{n_3:\ \text{fixing } n\})^{1/2}\,(N_1N_2N_3)^{-1}\,(\#\{n_2,n_3\})^{1/2}\\ &\lesssim N_1^{Cs_0}(N_1N_2N_3)^{-1}N_2N_3 \lesssim N_1^{-1+Cs_0}\end{aligned} \tag{8.86}$$

by trivially counting n_2, n_3.

The proof of Case (k) and (l) is hence complete, so is the proof of Proposition 2.1.

Appendix A: Time localization of $X^{s,b}$

In this section, we summarize several standard time localization facts for the $X^{s,b}$ space, and also briefly recall the proof of Lemma 3.5. The presentation mainly follows that from [6]. Here ϕ is a fixed time cut off function. There are several basic facts about the $X^{s,b}$ space that we can recall below. We have

$$\|\phi(t/\delta)u\|_{X^{s,b}} \lesssim_b \|u\|_{X^{s,b}}, \quad 0 < b < \frac{1}{2} \tag{A.1}$$

$$\|\phi(t/\delta)u\|_{X^{s,b}} \lesssim_b \delta^{\frac{1-2b}{2}}\|u\|_{X^{s,b}} \quad \frac{1}{2} < b < 1. \tag{A.2}$$

Also, Hausdorff–Young inequality gives the following estimate which is useful in the interpolation

$$\|\phi(t)u\|_{L^4_{t,x}} \lesssim_\epsilon \|u\|_{X^{1/2,\frac{1}{4}+\epsilon}}, \tag{A.3}$$

which can be compared to estimates (95), (96) on page 26 of [6].

In what follows, one should think $1 \gg s_p \gg \epsilon > 0$. We will only do proof for (3.12) in Lemma 3.5.

Via Strichartz estimate and interpolation of Hausdorff Young inequality, one can obtain

$$\|\phi(t)u\|_{L^4_{t,x}} \lesssim_\epsilon \|u\|_{X^{3\epsilon,\frac{1}{2}-\epsilon}} \tag{A.4}$$

(One may change the 3 in the above to any number larger than 2.) Similarly, for $p > 4$, one can obtain

$$\|\phi(t)u\|_{L^p_{t,x}} \lesssim_\epsilon \|u\|_{X^{s_p+10\epsilon,\epsilon}} \tag{A.5}$$

There are the following two Hölder inequalities,

(1)

$$\|\phi(t/\delta)u\|_{L^4_{t,x}} \le \|\phi(t/\delta)u\|^{\theta_p}_{L^2_{t,x}}\|\phi(t)u\|^{1-\theta_p}_{L^p_{t,x}}, \tag{A.6}$$

where one has $\frac{1}{4} = \frac{\theta_p}{2} + \frac{1-\theta_p}{p}$. $\theta_p = \frac{s_p}{1+s_p} \ge \frac{1}{2}s_p$

(2)

$$\|\phi(t/\delta)u\|_{L^2_{t,x}} \le \delta^{1/4}\|\phi(t)u\|_{L^4_{t,x}} \tag{A.7}$$

One derives

$$\begin{aligned}
&\|\phi(t/\delta)u\|_{L^4_{t,x}} \\
&\le \|\phi(t\delta)u\|^{\theta_p}_{L^2_{t,x}} \|\phi(t)u\|^{1-\theta_p}_{L^p_{t,x}} \\
&\le \|\phi(t\delta)u\|^{\theta_p}_{L^4_{t,x}} \delta^{\theta_p/4} \|\phi(t)u\|^{1-\theta_p}_{L^p_{t,x}} \\
&\lesssim_\epsilon \|u\|^{\theta_p}_{X^{3\epsilon,\frac{1}{2}-\epsilon}} \|u\|^{1-\theta_p}_{X^{sp+10\epsilon,\frac{1}{2}-\epsilon}} \delta^{\frac{\theta_p}{4}}
\end{aligned} \tag{A.8}$$

Note that

$$\theta_p = \frac{s_p}{1+s_p} \ge \frac{1}{2} s_p. \tag{A.9}$$

Thus, to summarize, when $s \ll 1$, and $\epsilon \ll s$, one has

$$\|\phi(t/\delta)u\|_{L^4_{t,x}} \lesssim_\epsilon \|u\|_{X^{s+10\epsilon,\frac{1}{2}-\epsilon}} \delta^{\frac{s}{4}}, \tag{A.10}$$

which, for convenience, can be written as

$$\|\phi(t/\delta)u\|_{L^4_{t,x}} \lesssim_\epsilon \|u\|_{X^{s,\frac{1}{2}-\epsilon}} \delta^{s/8}. \tag{A.11}$$

Localizing at frequency N, this gives Lemma 3.5 for balls B of radius N, which are centered at origin point. To prove general B centered at n_0, one simply observes

$$\sum_{n\in B} a_n e^{inx} e^{in^2 t} e^{i\lambda t} = \sum_{|n-n_0|\le N} a_n e^{i(n-n_0)(x+2n_0)} e^{i(n-n_0)^2 t} e^{i\lambda t} e^{in_0 x} e^{-in_0^2 t} \tag{A.12}$$

and the $L^4_{t,x}$ norm of a function is invariant under multiplying $e^{in_0 x} e^{-in_0^2 t}$ and doing space translation in x variable by n_0. This ends the proof.

Appendix B: Proof of Lemma 3.6, 3.7, 3.8

We briefly sketch the proof of those three Lemmata here.

We start with Lemma 3.8. Let $h(n,t)$, $f_i(n,t)$ be space of h, f_i, and we will also short handedly write them as $h(n)$, $f_i(n)$. We only prove

$$\begin{aligned}
&\left| \int \phi(t/\delta) h \mathcal{N}_2(P_{N_1} f_1, P_{N_1} f_2, P_{N_1} f_3) \right| \\
&\lesssim (\delta^{1/4} \|P_{N_1} h\|_{X_{0,1-b_0}} \|f_1\|_{X^{0,b_0}} \sup_{|n|\sim N_1} \Pi_{j\ne 1} \|f_j(n) e^{inx}\|_{X_{0,b_0}}).
\end{aligned} \tag{B.1}$$

To see this, observe

$$\left|\int \phi(t/\delta) h \mathcal{N}_2(P_{N_1} f_1, P_{N_1} f_2, P_{N_1} f_3)\right| = \left|\sum_{|n|\sim N_1} \int \phi(t/\delta)\bar{h}(n) f_1(n) \bar{f}_2(n) f_3(n)\right|$$
$$\lesssim \|\phi(t/\delta)\|_{L_t^2} \|\phi(t) h(n)\|_{L_t^2} \|\phi(t) f_1\|_{L_t^\infty} \sup_{|n|\sim N_1} \Pi_{j\neq 1} \|f_j\|_{L_t^\infty}. \tag{B.2}$$

Now we have, (note that one only has one mode in all the estimates below)

$$\|\phi(t)h(n)\|_{L_t^2} \lesssim \|h(n)e^{inx}\|_{X^{0,1-b_0}}, \|f_i(n)\|_{L_t^\infty} \lesssim \|f_i(n)e^{inx}\|_{X^{0,b_0}} \tag{B.3}$$

then, (B.1) will follow from (B.2) by Cauchy Schwarz.

We turn to Lemma 3.6. We start with (3.14) to (3.17). Estimates (3.14), (3.16) follows from (3.11), (3.12) via Hölder inequality. We point out that the naive loss will be $N_1^{C\epsilon}$ rather than $\max(N_2, N_3)^{C\epsilon}$, but this can be handled by a standard L^2 orthogonality argument, See, for example,[2,6] for more details. Now we show how to derive (3.17) from (3.16). We shall see that (3.15) can be derived similarly form (3.14).

Recall that we used the notation

$$f_i(x,t) = \sum_n f_i(n,t) e^{inx}, \quad i = 1, 2.3 \tag{B.4}$$

i.e. $f_i(n,t)$ is the space Fourier transform. For the sake of convenience, we denote $f_i(n,t)$ with $f_i(n)$. Similarly, we wrtie $h = \sum h(n,t)e^{inx}$.

Given (3.16), in order to derive (3.17), we need to further prove

- If $N_1 \sim N_2$

$$\sum_{n_1\sim|N_1|, n_3\sim N_3} \left|\int \phi(t/\delta \bar{h}(n_1) f_1(n_1) \bar{f}_2(n_3) f_3(n_3)\right|$$
$$\lesssim \delta^{1/10} (\max(N_2, N_3))^{C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}} \tag{B.5}$$

- If $N_2 \sim N_3$

$$\sum_{n_1\sim N_1, n_2\sim N_2} \left|\int \phi(t/\delta) \bar{h}(n_1) f_1(n_1) \bar{f}_2(n_2) f_3(n_2)\right|$$
$$\lesssim \delta^{1/10} (\max(N_2, N_3))^{C\epsilon_0} \|h\|_{X^{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}} \tag{B.6}$$

- If $N_1 \sim N_2 \sim N_3$,

$$\int |\phi(t/\delta) h \mathcal{N}_2(P_{N_1} f_1, P_{N_1} f_2, P_{N_1} f_3)|$$
$$\lesssim \delta^{1/10} (\max(N_2, N_3))^{C\epsilon_0} \|h\|_{X_{0,1-b_0}} \prod_i \|P_{N_i} f_i\|_{X^{0,b_0}} \tag{B.7}$$

Estimate (B.7) follows from Lemma 3.8. The proof of estimates (B.5) and (B.6) are similar, and we only work on (B.5). Note that the integration on the left side is only in t. One has, (by Sobolev embedding in the t variable if necessary), that

$$\begin{aligned}
&\|h(n_1)\|_{L_t^2} \lesssim \|h(n_1)e^{in_1x}\|_{X^{0,1-b_0}},\\
&\|f(n_1)\|_{L_t^\infty} \lesssim \|f(n_1)e^{in_1x}\|_{X^{0,b_0}},\\
&\|f_i(n_3)\|_{L_t^\infty} \lesssim \|f_i(n_3)e^{in_3x}\|_{X^{0,b_0}}.
\end{aligned} \tag{B.8}$$

Then the desired estimates follow from a Hölder inequality in t and Cauchy Schwarz inequality in n_1, n_3. Estimates (3.18), (3.20), (3.21) are of similar flavor. We prove (3.18) and leave the rest to the interested readers. Estimate (3.18) follows from the following four estimates.

- $$\begin{aligned}
&\left|\int \psi(t)\bar{h}P_{N_1}f_1P_{N_2}\bar{f}_2P_{N_3}f_3\right|\\
&\lesssim (\max(N_2,N_3))^{C\epsilon_0}|h\|_{X^{0,1/3}}\|\sup_J\|P_Jf_1\|_{L_{t,x}^\infty}\|f_2\|_{X^{0,1/3}}\|f_3\|_{X^{0,1/3}},
\end{aligned} \tag{B.9}$$

- If $N_1 \sim N_2$,

$$\begin{aligned}
&\sum_{n_1\sim|N_1|,n_3\sim N_3}\left|\int \psi(t)\overline{h(n_1)}f_1(n_1)\bar{f}_2(n_3)f_3(n_3)\right|\\
&\lesssim \|P_{N_1}f_1\|_{X^{0,b_0}}\|P_{N_1}f_2\|_{X^{0,1/3}}\|P_{N_3}f_3\|_{X^{0,1/3}}\|P_{N_3}h\|_{X^{0,1/3}},
\end{aligned} \tag{B.10}$$

- If $N_2 \sim N_3$,

$$\begin{aligned}
&\sum_{n_1\sim N_1,n_2\sim N_2}\left|\int \psi(t)\overline{h(n_1)}f_1(n_1)\bar{f}_2(n_2)f_3(n_3)\right|\\
&\lesssim \|P_{N_1}f\|_{X^{0,b_0}}\|P_{N_2}f_2\|_{X^{0,1/3}}\|P_{N_2}f_3\|_{X^{0,1/3}}\|P_{N_1}h\|_{X^{0,1/3}}
\end{aligned} \tag{B.11}$$

- If $N_1 \sim N_2 \sim N_3$

$$\begin{aligned}
&\left|\int \psi(t)\bar{h}\mathcal{N}_2(P_{N_1}f_1P_{N_2}\bar{f}_2P_{N_3}f_3)\right|\\
&\lesssim \min(\|P_{N_1}h\|_{X^{0,1-b_0}}\|f_i\|_{X^{0,b_0}}\sup_{|n|\sim N_1}\prod_{j\neq i}\|f_j(n)e^{inx}\|_{X^{0,b_0}}
\end{aligned} \tag{B.12}$$

Again estimate (B.12) follows from Lemma 3.8. We will only prove estimate (B.9), (B.10). The proof of (B.11) is similar to that for (B.10).

We start with (B.9). We may only consider the case $N_2 \geq N_3$, as the case $N_2 \leq N_3$ can be proved similarly.

We may further only consider the case $N_1 \gg N_2$, otherwise one may replace P_J by $P_{<N_1}$. Observe that (using L^2 orthogonality),

$$\begin{aligned}\int \psi(t)\bar{h} P_{N_1} f_1 P_{N_2} \bar{f}_2 P_{N_3} f_3 &= \sum_J \int \psi(t)\bar{h} P_J f_1 P_{N_2} \bar{f}_2 P_{N_3} f_3 \\ &= \sum_J \int \psi(t) P_J \bar{h} P_J f_1 P_{N_2} \bar{f} P_{N_3} f_3. \end{aligned} \tag{B.13}$$

For each J, we may estimate as follows,

$$\begin{aligned}&\left| \int\int \psi(t) P_J \bar{h} P_J f_1 P_{N_2} \bar{f} P_{N_3} f_3 \right| \\ &\lesssim \|\phi(t) P_J h\|_{L^3_{t,x}} \|\phi(t) P_J f_1\|_{L^\infty_{t,x}} \|\phi(t) P_{N_2} f_2\|_{L^3_{t,x}} \|\phi(t) P_{N_3} f_3\|_{L^3_{t,x}}, \end{aligned} \tag{B.14}$$

where without loss of generality assumed $\psi(t) = \phi(t)^4$ for some well localized $\phi(t)$.

Using Estimate (3.7) to control the L^3 norm in (B.14) and applying a Cauchy Schwarz in J, the desired estimate then follows.

Lemma 3.7 can be proved similarly as Lemma 3.6.

Appendix C: A Cauchy–Schwarz type inequality

We summarize a (deterministic) Cauchy–Schwarz type inequality, that is often used in random data type problems. For simplicity, let a_{ij}, b_j be real numbers, assume that

$$\sum_j b_j^2 \lesssim 1, \tag{C.1}$$

which of course implies

$$\sum_{j,j'} b_j^2 b_{j'}^2 \lesssim 1. \tag{C.2}$$

Then, we have

$$\sum_i |\sum_j a_{ij} b_j|^2 = \sum_i \sum_{j,j'} a_{ij} a_{ij'} b_j b_{j'} = \sum_i \sum_j a_{ij} a_{ij} b_j^2 + \sum_i \sum_{j \neq j'} a_{ij} a_{ij'} b_j b_{j'} \tag{C.3}$$

Note that

$$\sum_i |\sum_j a_{ij} a_{ij} b_j|^2 \lesssim \sup_j \sum_i a_{ij}^2 \tag{C.4}$$

and, by Cauchy inequality,

$$\sum_i \sum_{j\neq j'} a_{ij}a_{ij'}b_j b_{j'} = \sum_{j\neq j'} b_j b_{j'} \sum_i a_{ij}a_{ij'} \lesssim \left(\sum_{j,j'} b_j^2 b_{j'}^2\right)^{1/2} \left\{\sum_{j\neq j'} |\sum_i a_{ij}a_{ij}|^2\right\}^{1/2} \tag{C.5}$$

To summarize, and by simple generalization to the complex case, one has

Lemma C.1 *Assume* $\sum_j |b_j|^2 \lesssim 1$, *then*

$$\sum_i |\sum_j a_{ij}b_j|^2 \lesssim \max_j \sum_i |a_{ij}|^2 + \left(\sum_{j\neq j'} |\sum_i a_{ij}\bar{a}_{ij'}|^2\right)^{1/2} \tag{C.6}$$

One can also easily write down, via the dual estimate,

Lemma C.2 *Assume* $\sum_j |b_j|^2 \lesssim 1$, *then*

$$\sum_i |\sum_j a_{ij}b_j|^2 \lesssim \max_i \sum_j |a_{ij}|^2 + \left(\sum_{i\neq i'} |\sum_j a_{i'j}\bar{a}_{ij}|^2\right)^{1/2} \tag{C.7}$$

References

1. Bombieri, E., Pila, J., et al.: The number of integral points on arcs and ovals. Duke Math. J **59**(2), 337–357 (1989)
2. Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, part i, nonlinear Schrödinger Equations. Geom. Funct. Anal. **3**(2), 107–156 (1993)
3. Bourgain, J.: Periodic nonlinear Schrödinger equation and invariant measures. Commun. Math. Phys. **166**(1), 1–26 (1994)
4. Bourgain, J.: Invariant measures for the 2D-defocusing nonlinear Schrödinger equation. Commun. Math. Phys. **176**(2), 421–445 (1996)
5. Bourgain, J., Demeter, C.: The proof of the L^2 decoupling conjecture. Ann. Math. **182**, 351–389 (2015)
6. Caffarelli, L., Weinan, E.: Hyperbolic equations and frequency interactions, vol. 5. In: Hyperbolic Equations and Frequency Interactions (1999)
7. Deng, Y., Germain, P., Guth, L.: Strichartz estimates for the Schrödinger equation on irrational tori. J. Funct. Anal. **273**(9), 2846–2869 (2017)
8. Deng, Yu, Nahmod, Andrea, Yue, Haitian: Invariant Gibbs measure and global strong solutions for nonlinear Schrödinger equations in dimension two. arXiv:1910.08492 (2019)
9. Glimm, J., Jaffe, A.: Quantum Physics. Springer, New York (1989)
10. Lebowitz, J., Rose, H., Speer, E.: Statistical mechanics of the nonlinear Schrödinger equation. J. Stat. Phys. **50**, 657–687 (1988)
11. Thomann, L., Tzvetkov, N.: Gibbs measure for the periodic derivative nonlinear Schrödinger equation. Nonlinearity **23**(11), 2771 (2010)
12. Tzvetkov, N.: Construction of a Gibbs measure associated to the periodic Benjamin–Ono equation. Probab. Theory Relat. Fields **146**(3–4), 481 (2010)
13. Tzvetkov, N., Sun, C.: Concerning the pathological set in the context of probabilistic well-posedness. arXiv:2001.10293 (2020)