ANALYSIS & PDE

Volume 11 No. 4 2018

CHENJIE FAN, GIGLIOLA STAFFILANI, HONG WANG AND BOBBY WILSON

ON A BILINEAR STRICHARTZ ESTIMATE ON IRRATIONAL TORI

msp

ANALYSIS AND PDE
Vol. 11, No. 4, 2018
dx.doi.org/10.2140/apde.2018.11.919

ON A BILINEAR STRICHARTZ ESTIMATE ON IRRATIONAL TORI

CHENJIE FAN, GIGLIOLA STAFFILANI, HONG WANG AND BOBBY WILSON

We prove a bilinear Strichartz-type estimate for irrational tori via a decoupling-type argument, as used by Bourgain and Demeter (2015), recovering and generalizing a result of De Silva, Pavlović, Staffilani and Tzirakis (2007). As a corollary, we derive a global well-posedness result for the cubic defocusing NLS on two-dimensional irrational tori with data of infinite energy.

1. Introduction

Bourgain and Demeter [2015] proved the full range of Strichartz estimates for the Schrödinger equation on tori as a consequence of the ℓ^2 decoupling theorem. In this paper we prove in full generality the analog of the improved Strichartz estimate that first appeared in [De Silva et al. 2007] for rational tori.

1A. ***Statement of the problem and main results.*** Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus, and let $\alpha_1, \dots, \alpha_{d-1} \in \left[\frac{1}{2}, 1\right]$; we define d-dimensional torus $\mathbb{T}^d$ as $\mathbb{T}^d = \mathbb{T} \times \alpha_1 \mathbb{T} \times \cdots \times \alpha_{d-1}\mathbb{T}$. We say that the torus is irrational if at least one α_i is irrational. The torus is rational otherwise. For any $\lambda \geq 1$, we define $\mathbb{T}^d_\lambda$ as a rescaling of $\mathbb{T}^d$ by λ; i.e.,

$$\mathbb{T}^d_\lambda = \lambda \mathbb{T}^d = (\lambda\mathbb{T}) \times (\alpha_1 \lambda \mathbb{T}) \times \cdots \times (\alpha_{d-1}\lambda\mathbb{T}).$$

When $\lambda \to \infty$, one should think of $\mathbb{T}_\lambda$ as a large torus approximating $\mathbb{R}^d$. We consider the following Cauchy problem for the linear Schrödinger equation on $\mathbb{T}^d_\lambda$:

$$\begin{cases} iu_t - \Delta u = 0, & (t,x) \in \mathbb{R} \times \mathbb{T}^d_\lambda, \\ u(0,x) = u_0, & u_0 \in L^2(\mathbb{T}^d_\lambda). \end{cases} \tag{1-1}$$

Let $U_\lambda(t)u_0$ be the solution to (1-1), and let

$$\Lambda_\lambda := \frac{1}{\lambda}\left(\mathbb{Z} \times \frac{1}{\alpha_1}\mathbb{Z} \times \cdots \times \frac{1}{\alpha_{d-1}}\mathbb{Z}\right).$$

One has

$$U_\lambda(t)u_0(x) = \frac{1}{\lambda^{d/2}} \sum_{k \in \Lambda_\lambda} e^{2\pi k i x - |2\pi k|^2 i t}\hat{u}_0(k). \tag{1-2}$$

Our main theorem is the following bilinear refined Strichartz estimate.

Fan and Staffilani are partially supported by NSF DMS 1362509 and DMS 1462401.
MSC2010: 35Q55, 42B37.
Keywords: decoupling, bilinear Strichartz.

Theorem 1.1. *Let $\phi_1, \phi_2 \in L^2(\mathbb{T}_\lambda)$ be two initial data such that* $\operatorname{supp}\hat{\phi}_i \subset \{k : |k| \sim N_i\}$, $i = 1, 2$, *for some large $N_1 \geq N_2$, and let $\eta(t)$ be a time cut-off function*, $\operatorname{supp}\eta \subset [0, 1]$. *Then when $d = 2$,*

$$\|\eta(t)U_\lambda\phi_1 \cdot \eta(t)U_\lambda\phi_2\|_{L^2_{x,t}} \lesssim N_2^\epsilon \left(\frac{1}{\lambda} + \frac{N_2}{N_1}\right)^{1/2} \|\phi_1\|_{L^2}\|\phi_2\|_{L^2}, \tag{1-3}$$

and when $d \geq 3$

$$\|\eta(t)U_\lambda\phi_1 \cdot \eta(t)U_\lambda\phi_2\|_{L^2_{x,t}} \lesssim N_2^\epsilon \left(\frac{N_2^{d-3}}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2} \|\phi_1\|_{L^2}\|\phi_2\|_{L^2}. \tag{1-4}$$

We note that when $d = 2$, $N_1 = N_2$, and $\lambda = 1$, estimate (1-3) recovers the Strichartz inequality for the (irrational) torus after an application of Hölder's inequality, up to an N_2^ϵ-loss. When $\lambda \to \infty$, estimates (1-3) and (1-4) are consistent with the bilinear Strichartz inequality in $\mathbb{R}^{d+1}$ [Bourgain 1998]. Up to the N_2^ϵ-loss, inequality (1-3) is sharp.

Furthermore, when $\lambda \geq N_1$, the estimates fall into the so-called semiclassical regime in which the geometry of $\mathbb{T}_\lambda$ is irrelevant. We refer to [Hani 2012] for the same estimate (without N_2^ϵ-loss) on general compact manifolds. On the torus, our result improves the estimate in that paper for $\lambda \leq N_1$. Estimates (1-3) and (1-4) rely on the geometry of the torus and cannot hold on general compact manifolds.

Remark 1.2. It may also be interesting to consider trilinear estimates. In fact when one considers the quintic nonlinear Schrödinger equation, as in [Herr et al. 2011; Ionescu and Pausader 2012], trilinear estimates are fundamental. See also [Ramos 2016].

We will derive Theorem 1.1 from some bilinear decoupling-type estimates. We first introduce some basic notation.

Let P be the truncated paraboloid in $\mathbb{R}^{d+1}$,

$$P = \{(\xi, |\xi|^2) : \xi \in \mathbb{R}^d,\ |\xi| \lesssim 1\}. \tag{1-5}$$

For any function f supported on P, we define

$$Ef = \widehat{f d\sigma}, \tag{1-6}$$

where σ is the measure on P.

Note a function supported on P can be naturally understood as a function supported on the ball $B = \{\xi \in \mathbb{R}^d : |\xi| \lesssim 1\}$.

By a slight abuse of notation, for a function f supported in the ball B in $\mathbb{R}^d$, we also define

$$Ef(x, t) = \int_B e^{-2\pi i(\xi\cdot x + |\xi|^2 t)} f(\xi)\, d\xi. \tag{1-7}$$

One can see that the two definitions of Ef are essentially the same since P projects onto B.

We decompose P as a finitely overlapping union of caps θ of radius δ. Here a cap θ of radius δ is the set $\theta = \{\xi \in P : |\xi - \xi_0| \lesssim \delta\}$ for some fixed $\xi_0 \in P$. We define $Ef_\theta = \widehat{f_\theta d\sigma}$, where f_θ is f restricted to θ. We use a similar definition also when f is a function supported on the unit ball in $\mathbb{R}^d$. We have $Ef = \sum_\theta Ef_\theta$.

Now, we are ready to state our main decoupling-type estimate.

Theorem 1.3. *Given $\lambda \geq 1$, $N_1 \geq N_2 \geq 1$, let f_1 be supported on P where $|\xi| \sim 1$, and let f_2 be supported on P where $|\xi| \sim N_2/N_1$. Let $\Omega = \{(t, x) \in [0, N_1^2] \times [0, (\lambda N_1)^2]^d\}$. For a finitely overlapping covering of the ball $B = \{|\xi| \leq 1\}$ of caps $\{\theta\}$, $|\theta| = 1/(\lambda N_1)$, we have the following estimate. For any small $\epsilon > 0$, when $d = 2$,*

$$\|Ef_1 Ef_2\|_{L^2_{\mathrm{avg}}(w_\Omega)} \lesssim_\epsilon (N_2)^\epsilon \lambda^{d/2} \left(\frac{1}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2} \prod_{j=1}^{2} \left(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_\Omega)}\right)^{1/2}, \tag{1-8}$$

and when $d \geq 3$,

$$\|Ef_1 Ef_2\|_{L^2_{\mathrm{avg}}(w_\Omega)} \lesssim_\epsilon (N_2)^\epsilon \lambda^{d/2} \left(\frac{N_2^{d-3}}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2} \prod_{j=1}^{2} \left(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_\Omega)}\right)^{1/2}, \tag{1-9}$$

where w_Ω is a weight adapted to Ω.

The presence of the weight w in these estimates is standard. We list the basic properties of w in Section 1D, and one can refer to [Bourgain and Demeter 2017] for more details. The notation $L_{\mathrm{avg}}(w_\Omega)^2$ is explained in Section 1C.

The proof of Theorem 1.3 gives another proof of the linear decoupling theorem in [Bourgain and Demeter 2015] in dimension $d = 2$, and does not rely on multilinear Kakeya or multilinear restriction theorems in $\mathbb{R}^3$. The proof of Theorem 1.3 in dimension $d \geq 3$ relies instead on linear decoupling in $\mathbb{R}^{d+1}$ [Bourgain and Demeter 2015].

Remark 1.4. The estimates in Theorems 1.1 and 1.3 are sharp up to an N_2^ϵ. See the Appendix for examples.

Remark 1.5. The N_2^ϵ-loss in Theorem 1.1 is typical if one wants to directly use a decoupling-type argument. It may be possible to remove N_2^ϵ in the mass supercritical setting (in our case, this means $d \geq 3$), using the approach in [Killip and Vişan 2016], where the scale-invariant Strichartz estimates are studied.

Remark 1.6. Similar bilinear estimates for dimension $d \geq 3$ were also considered in [Killip and Vişan 2016] for nonrescaled tori; see Lemma 3.3 in that paper. On the other hand in this work we also consider the $d = 2$ case, which is mass critical.

1B. ***Background and motivation.*** System (1-1) and the bilinear estimates (1-3) and (1-4) naturally appear in the study of the following nonlinear Schrödinger equation on the *nonrescaled* tori:

$$\begin{cases} iu_t + \Delta u = |u|^2 u, \\ u(0) = u_0 \in H^s(\mathbb{T}^d). \end{cases} \tag{1-10}$$

Let us focus for a moment on the $d = 2$ case. The Cauchy problem is said to be locally well-posed in $H^s(\mathbb{T}^d)$ if for any initial data $u_0 \in H^s(\mathbb{T}^d)$ there exists a time $T = T(\|u_0\|_s)$ such that a unique solution to the initial value problem exists on the time interval $[0, T]$. We also require that the data-to-solution map is continuous from $H^s(\mathbb{T}^d)$ to $C_t^0 H_x^s([0, T] \times \mathbb{T}^d)$. If $T = \infty$, we say that a Cauchy problem is globally well-posed.

The initial value problem (1-10) is locally well-posed for initial data $u_0 \in H^s$, $s > 0$, via Strichartz estimates. Note that using iteration, by the energy conservation law, i.e.,

$$E(u(t)) = E(u_0) = \frac{1}{2}\int |\nabla u|^2 + \frac{1}{4}\int |u|^4,$$

all initial data in $H^1(\mathbb{T}^2)$ give rise to a global solution. Next, by the nowadays standard I-method [Colliander et al. 2002] by considering a modified version of the energy, in the rational torus case, it was proved in [De Silva et al. 2007] that (1-10) is indeed globally well-posed for initial data in H^s, $s > \frac{2}{3}$. The key estimate there was in fact (1-3) for linear solutions on *rescaled* tori, which we prove here to be available also for irrational tori.

The proof for (1-3) presented in [De Silva et al. 2007] is only for rational tori since it relies on certain types of counting lemmata that cannot directly work on irrational tori. One of the main purposes of this work in fact is to extend results on rational tori to irrational ones.

Based on the discussion we just made, as a corollary of Theorem 1.1, we have:

Corollary 1.7. *The initial value problem* (1-10) *defined on any torus* $\mathbb{T}^2$ *is globally well-posed for initial data in* $H^s(\mathbb{T}^2)$ *with* $s > \frac{2}{3}$.

Remark 1.8. Results such as Corollary 1.7 usually also give a control on the growth of Sobolev norms of the global solutions. We do not address this particular question here. We instead refer the reader to the recent work [Deng and Germain 2017].

The original Strichartz estimates needed to prove the local well-posedness of Cauchy problems such as (1-10) were first obtained in [Bourgain 1993] via number-theoretical-related counting arguments for rational tori. Recently, the striking proof of the ℓ^2 decoupling theorem [Bourgain and Demeter 2015] provided a completely different approach from which all the desired Strichartz estimates on tori, both rational and irrational, follow. This approach in particular does not depend on counting lattice points. See also [Guo et al. 2014; Deng et al. 2017]. The method of proof we implement in this present work is mostly inspired by [Bourgain and Demeter 2015] and the techniques used to prove the ℓ^2 decoupling theorem.

We quickly recall the main result in [Bourgain and Demeter 2015]. Let P be a unit parabola in $\mathbb{R}^{d+1}$, covered by finitely overlapping caps θ of radius $1/R$. Let f be a function defined on P; then one has for any $\epsilon > 0$ small,

$$\|Ef\|_{L^p(w_{B_{R^2}})} \lesssim_\epsilon R^\epsilon (R^2)^{d/4-(d+2)/(2p)} \left(\sum_\theta \|Ef_\theta\|^2_{L^p(B_{R^2})}\right)^{1/2}, \quad p \geq \frac{2(d+2)}{d}. \tag{1-11}$$

Note that (1-11) corresponds to Theorem 1.1 in [Bourgain and Demeter 2015], and the dimension n in the estimate (2) there corresponds to our $d+1$. Also note that the linear decoupling (1-11) not only works for those f exactly supported on P, but those f supported in an R^{-2} neighborhood of P, and in this case, cap θ would be replaced by the R^{-2} neighborhood of the original θ; see Theorem 1.1 in [Bourgain and Demeter 2015].

We remark that one key feature of this decoupling-type estimate is that one needs to work on a larger scale in physical space, i.e., the scale R^2 rather than R, in order to observe the decoupling phenomena. The proper observational scale dictated by Heisenberg's uncertainty principle is R.

Indeed, one principle, which is usually called *parallel decoupling*, indicates that if decoupling happens in a small region, then decoupling happens in a large region as well. We state a bilinear version of parallel decoupling below.

Lemma 1.9 [Bourgain and Demeter 2015; 2017]. *Let D be a domain, and $D = D_1 \cup D_2 \cup \cdots \cup D_J$, $D_i \cap D_j = \varnothing$. If for some constant $A > 0$ and for functions h_1, h_2, defined on the unit parabola, one has*

$$\|Eh_1 Eh_2\|_{L^2_{\mathrm{avg}}(w_{D_i})} \le A \prod_{j=1}^{2} \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Eh_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{D_i})} \Bigr)^{1/2}, \quad i = 1, \dots, J, \tag{1-12}$$

then one also has

$$\|Eh_1 Eh_2\|_{L^2_{\mathrm{avg}}(w_D)} \le A \prod_{j=1}^{2} \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Eh_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_D)} \Bigr)^{1/2}. \tag{1-13}$$

The proof of this particular formulation of parallel decoupling follows by Minkowski's inequality.

As it exists, parallel decoupling is a principle rather than a concrete lemma. We state the version here solely for concreteness. It should be easy to generalize the lemma under different conditions.

1C. *Notation.* We write $A \lesssim B$ if $A \le CB$ for a constant $C > 0$, and $A \sim B$ if both $A \lesssim B$ and $B \lesssim A$. We say $A \lesssim_\epsilon B$ if the constant C depends on ϵ. Similarly for $A \sim_\epsilon B$. For a Borel set, $E \subset \mathbb{R}^d$, we denote the diameter of E by $|E|$ and the Lebesgue measure of E by $m(E)$.

We will use the usual function space L^p. We also use a (weighted) average version of L^p space; i.e.,

$$\|g\|_{L^p_{\mathrm{avg}}(A)} = \Bigl(\fint_A |g|^p \Bigr)^{1/p} := \Bigl(\frac{1}{m(A)} \int_A |g|^p \Bigr)^{1/p},$$

$$\|g\|_{L^p_{\mathrm{avg}}(w_A)} = \Bigl(\frac{1}{m(A)} \int |g|^p w_A \Bigr)^{1/p},$$

where w_A is a weight function described below.

For any function f, we use $\hat{f}$ to denote its Fourier transform. When we say *unit ball*, we refer to a ball of radius $r \sim 1$. We will often identify a torus as a bounded domain in Euclidean space; for example, we will view $(\mathbb{R}/\mathbb{Z})^d$ as $[0,1]^d \subset \mathbb{R}^d$. In this work, Ω is used to denote the domain $[0, N_1^2] \times [0, (\lambda N_1)^2]^d \subset \mathbb{R}^{d+1}$.

1D. *The weight w_A.* If h is a Schwartz function whose Fourier transform, $\hat{h}$, is supported in a ball of radius $1/R$, we expect h to be essentially constant on balls of radius R, and essentially

$$\|h\|_{L^p_{\mathrm{avg}}(B_R)} \sim \|h\|_{L^2_{\mathrm{avg}}(B_R)} \sim \|h\|_{L^\infty(B_R)}. \tag{1-14}$$

Expression (1-14) is not rigorous, and the introduction of the weight w_{B_R} is a standard way to overcome this technical difficulty. We refer to Lemma 4.1 in [Bourgain and Demeter 2017] for a more detailed discussion of the weight function.

For any bounded open convex set A, the weight function w_A might change from line to line and from the left-hand side of the inequality to the right-hand side, and satisfies the properties:

- $\int w_A \sim m(A)$.
- $w_A \gtrsim 1$ on A, and rapidly (polynomial-type) decays outside A.

We will usually define A to be a ball, or the product of balls in this paper.

Furthermore, let B_R be a ball centered at 0, and let μ_{B_R} be a function such that $\widehat{\mu_{B_R}}$ is about $1/(m(B_{1/R}))$ on $B_{1/R}$, and supported in $B_{2/R}$. Then μ_{B_R} is about 1 on B_R, and decays faster than any polynomial outside of B_R. Additionally, $\mu_{B_R}^2$ is positive, decays faster than any polynomial outside of B_R and is Fourier-supported in $B_{4/R}$. We take translations B' of B_R to cover the whole space, and we denote by $\mu_{B'}$ the corresponding translation of μ_{B_R} and $w_{B_R}(B') = \max_{x\in B'} w_{B_R}$. We have the useful property,

$$w_{B_R}(x) \le \sum_{B'} w_{B_R}(B') 1_{B'}(x) \lesssim \sum_{B'} w_{B_R}(B') \mu_{B'}^2(x) \lesssim w_{B_R}(x). \tag{1-15}$$

The last inequality follows from the fact that $\mu_{B'}^2$ decays faster than any polynomial outside of B'.

Lemma 1.10. *For a function f supported in $B_{1/R}$, for any $p<\infty$,*

$$\|Ef\|_{L^\infty(B_R)} \lesssim \|Ef\|_{L^p_{\mathrm{avg}}(\mu_{B_R})}.$$

We refer to the proof of Corollary 4.3 in [Bourgain and Demeter 2017] with the weight on the left-hand side being 1_{B_R} so that on the right-hand side we have a fast decay weight.

Remark 1.11. In general, Lemma 1.10 should hold for any convex set A and the dual convex body A^*.

2. Proof of Theorem 1.1 assuming Theorem 1.3

Assume Theorem 1.3, and let us prove Theorem 1.1. The argument below comes from the proof of discrete restriction and the Strichartz estimate on irrational tori assuming the ℓ^2 decoupling estimate; see Theorems 2.2 and 2.3 in [Bourgain and Demeter 2015]. The argument originally comes as observation due to Bourgain [2013]. We record it here for completeness.

Let ϕ_1, ϕ_2 be as in Theorem 1.1. We rescale ϕ_1 to be supported in the unit ball and rescale ϕ_2 to be supported in a ball of radius $\sim N_2/N_1$. Recall,

$$U_\lambda(t)\phi_j(x,t) = \frac{1}{\lambda^{d/2}} \sum_{\substack{k\in\Lambda_\lambda \\ k\sim N_1}} e^{2\pi i k\cdot x - |2\pi k|^2 t} \hat{\phi}_j(k). \tag{2-1}$$

We perform a change of variables $\xi = k/N_1$ and we let

$$h_j(\tau) = \frac{1}{\lambda^{d/2}} \sum_{\substack{\xi\in\Lambda_{\lambda N_1} \\ |\xi|\sim 1}} \hat{\phi}_j(\xi N_1)\delta_\xi(\tau), \quad j=1,2. \tag{2-2}$$

Note one can directly check that

$$U_\lambda(t)\phi_j(x,t) = Eh_j(-2\pi N_1 x, (2\pi)^2 N_1^2 t). \tag{2-3}$$

Without loss of generality, we suppress the constants -2π and $(2\pi)^2$.

Let $Q_0=[0,N_1^2]\times\mathbb{T}^d_{\lambda N_1}$ and let us view $\mathbb{T}^d_{\lambda N_1}$ as a compact set in $\mathbb{R}^d$. In particular, one can construct the associated weight function w_{Q_0}. Direct computation (via change of variables) gives

$$\|U_\lambda(t)\phi_1U_\lambda(t)\phi_2\|_{L^2([0,1]\times\mathbb{T}^d_\lambda)}\sim N_1^{-(d+2)/2}m(Q_0)^{1/2}\|Eh_1Eh_2\|_{L^2_{\text{avg}}(Q_0)} \tag{2-4}$$

and due to the periodicity of $Eh_i,\ i=1,2$, one has

$$\|Eh_1Eh_2\|_{L^2_{\text{avg}}(\Omega)}=\|Eh_1Eh_2\|_{L^2_{\text{avg}}(Q_0)}. \tag{2-5}$$

For a covering $\{\theta\}$ of caps of radius $1/(\lambda N_1)$, each cap θ contains at most one $\xi_\theta\in\Lambda_{\lambda N_1}$, corresponding to $k_\theta=N_1\xi_\theta\in\Lambda_\lambda$. Then

$$\|Eh_{j,\theta}\|_{L^4_{\text{avg}}(w_{Q_0})}\sim h_j(\xi_\theta)\sim\frac{1}{\lambda^d}\hat{\phi}_j(k_\theta),$$

and

$$\prod_{j=1}^2\biggl(\sum_{|\theta|=1/(\lambda N_1)}\|Eh_j\|^2_{L^4_{\text{avg}}(w_{Q_0})}\biggr)^{1/2}\sim\lambda^{-d}\prod_{j=1}^2\biggl(\frac{1}{\lambda^d}\sum_{k\in\Lambda_\lambda}|\hat{\phi}_j(k)|^2\biggr)^{1/2}\sim\lambda^{-d}\|\phi_1\|_{L^2}\|\phi_2\|_{L^2}.$$

For convenience of notation let

$$D_{\lambda,N_1,N_2}:=\begin{cases}1/\lambda+N_2/N_1 & \text{when } d=2,\\ N_2^{d-3}/\lambda+N_2^{d-1}/N_1 & \text{when } d\geq 3.\end{cases} \tag{2-6}$$

Recall that $\Omega=[0,N_1]^2\times[0,(\lambda N_1)^2]^d$; we apply Theorem 1.3 with $f_j=h_j$, and we have

$$\|Eh_1Eh_2\|_{L^2_{\text{avg}}(w_\Omega)}\lesssim_\epsilon (N_2)^\epsilon\lambda^{d/2}D^{1/2}_{\lambda,N_1,N_2}\prod_{j=1}^2\biggl(\sum_{|\theta|=1/(\lambda N_1)}\|Eh_{j,\theta}\|^2_{L^4_{\text{avg}}(w_\Omega)}\biggr)^{1/2}. \tag{2-7}$$

Note that Ω can be covered by Q such that $\{Q\}$ are finitely overlapping and each Q is a translation of Q_0. Since Eh_j are periodic on x, estimate (2-7) is equivalent to

$$\|Eh_1Eh_2\|_{L^2_{\text{avg}}(w_{Q_0})}\lesssim_\epsilon (N_2)^\epsilon\lambda^{d/2}D^{1/2}_{\lambda,N_1,N_2}\prod_{j=1}^2\biggl(\sum_{|\theta|=1/(\lambda N_1)}\|Eh_{j,\theta}\|^2_{L^4(w_{Q_0})}\biggr)^{1/2}. \tag{2-8}$$

Plugging (2-8) into (2-4) gives

$$\begin{aligned}\|U_\lambda(t)\phi_1U_\lambda(t)\phi_2\|_{L^2([0,1]\times\mathbb{T}^d_\lambda)}&\lesssim N_1^{-(d+2)/2}\cdot N_1m(\mathbb{T}^d_{\lambda N_1})^{1/2}\lambda^{-d}\cdot(N_2)^\epsilon\lambda^{d/2}D^{1/2}_{\lambda,N_1,N_2}\|\phi_1\|_{L^2}\|\phi_2\|_{L^2}\\ &\sim (N_2)^\epsilon D^{1/2}_{\lambda,N_1,N_2}\|\phi_1\|_{L^2}\|\phi_2\|_{L^2}\end{aligned}$$

and Theorem 1.1 follows.

The rest of the paper details the proof of Theorem 1.3.

3. An overview of the proof of Theorem 1.3

First, we reduce the proof of Theorem 1.3 to the following proposition.

Proposition 3.1. *Let τ_1 be a cap of radius N_2/N_1 supported at ξ with $|\xi| \sim 1$. Let τ_2 be a cap of radius N_2/N_1 supported at ξ with $|\xi| \sim N_2/N_1$. Let f_j be a function supported in τ_j. Then for any small $\epsilon > 0$, when $d = 2$*

$$\|Ef_1 Ef_2\|_{L^2_{\text{avg}}(w_\Omega)} \lesssim_\epsilon (N_2)^\epsilon \lambda^{d/2} \left(\frac{1}{\lambda} + \frac{N_2}{N_1}\right)^{1/2} \prod_{j=1}^{2} \Bigg(\sum_{\substack{|\theta|=1/(\lambda N_1)\\ \theta\subset\tau_j}} \|Ef_{j,\theta}\|^2_{L^4_{\text{avg}}(w_\Omega)} \Bigg)^{1/2}, \tag{3-1}$$

and when $d \geq 3$,

$$\|Ef_1 Ef_2\|_{L^2_{\text{avg}}(w_\Omega)} \lesssim_\epsilon (N_2)^\epsilon \lambda^{d/2} \left(\frac{N_2^{d-3}}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2} \prod_{j=1}^{2} \Bigg(\sum_{\substack{|\theta|=1/(\lambda N_1)\\ \theta\subset\tau_j}} \|Ef_{j,\theta}\|^2_{L^4_{\text{avg}}(w_\Omega)} \Bigg)^{1/2}. \tag{3-2}$$

Now, let f_1, f_2 be as in Proposition 3.1. We define $K_0(\lambda, N_1, N_2)$ to be the best constant such that

$$\|Ef_1 Ef_2\|_{L^2_{\text{avg}}(w_\Omega)} \leq \lambda^{d/2} K_0(\lambda, N_1, N_2) \prod_{j=1}^{2} \Bigg(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\text{avg}}(w_\Omega)} \Bigg)^{1/2}. \tag{3-3}$$

We also let $\widetilde{K}(\lambda, N_1, N_2)$ and $K(\lambda, N_1, N_2)$ be defined as the best constants such that

$$\|Ef_1 Ef_2\|_{L^2_{\text{avg}}(w_{[0,N_1^2]\times[0,\lambda N_1]^d})} \leq \lambda^{d/2} \widetilde{K}(\lambda, N_1, N_2) \prod_{j=1}^{2} \Bigg(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\text{avg}}(w_{[0,N_1^2]\times[0,\lambda N_1]^d})} \Bigg)^{1/2}, \tag{3-4}$$

$$\|Ef_1 Ef_2\|_{L^2_{\text{avg}}(w_{B_{N_1^2}})} \leq \lambda^{d/2} K(\lambda, N_1, N_2) \prod_{j=1}^{2} \Bigg(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\text{avg}}(w_{B_{N_1^2}})} \Bigg)^{1/2}. \tag{3-5}$$

Below we will prove that

$$\begin{aligned} K_0(\lambda, N_1, N_2) &\lesssim N_2^\epsilon \left(\frac{1}{\lambda} + \frac{N_2}{N_1}\right)^{1/2}, && d = 2,\\ K_0(\lambda, N_1, N_2) &\lesssim N_2^\epsilon \left(\frac{N_2^{d-3}}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2}, && d \geq 3. \end{aligned} \tag{3-6}$$

We point out here that by parallel decoupling, Lemma 1.9, one always has

$$\begin{aligned} K_0(\lambda, N_1, N_2) &\lesssim K(\lambda, N_1, N_2),\\ K_0(\lambda, N_1, N_2) &\lesssim \widetilde{K}(\lambda, N_1, N_2). \end{aligned} \tag{3-7}$$

The proof of Proposition 3.1 or equivalently (3-6) proceeds as follows. We first show:

Lemma 3.2. *When $\lambda \geq N_1$,*

$$\widetilde{K}(\lambda, N_1, N_2) \lesssim N_2^{\epsilon} \frac{N_2^{(d-1)/2}}{N_1^{1/2}}. \tag{3-8}$$

Note that when $\lambda \geq N_1$, Proposition 3.1 follows from (3-7) and Lemma 3.2.

Then, we show:

Lemma 3.3. *When $\lambda \leq N_1$,*

$$\begin{aligned} K(\lambda, N_1, N_2) &\lesssim N_2^{\epsilon} \left(\frac{1}{\lambda} + \frac{N_2}{N_1}\right)^{1/2}, && d = 2, \\ K(\lambda, N_1, N_2) &\lesssim N_2^{\epsilon} \left(\frac{N_2^{d-3}}{\lambda} + \frac{N_2^{d-1}}{N_1}\right)^{1/2}, && d = 3. \end{aligned} \tag{3-9}$$

From (3-7), clearly Proposition 3.1 follows from Lemmas 3.2 and 3.3.

The proof of Lemma 3.3 in dimension $d = 2$ relies on induction (of scale N_2). The proof of Lemma 3.3 in dimension in $d \geq 3$ is easier and more straightforward, (in some sense, it also relies on induction, but it is enough to induct only once.)

We first show the base case:

Lemma 3.4. *When $\lambda \leq N_1$ and $N_2 \lesssim 1$, we have $K(\lambda, N_1, N_2) \lesssim 1/\lambda^{1/2}$.*

Lemma 3.4 is not as useful in dimension $d \geq 3$, we indeed have a better estimate:

Lemma 3.5. *When $d \geq 3$, $\lambda \leq N_1$ and $\lambda \leq N_1/N_2^2$, we have $K(\lambda, N_1, N_2) \lesssim (N_2^{d-3}/\lambda)^{1/2}$.*

We then show the following lemma, which ensures that we only need to induct until $\lambda \leq N_1/N_2$, when $d = 2$, and until N_1/N_2 when $d \geq 3$.

Lemma 3.6. *Let $\lambda \leq N_1$.*

Let $d = 2$. Assume we have $K(\lambda, N_1, N_2) \leq \lambda^{-1/2}$ when $\lambda < N_1/N_2$. Then

$$K(\lambda, N_1, N_2) \leq N_2^{\epsilon} \frac{N_2^{(d-1)/2}}{N_1^{1/2}} \quad \text{when } \lambda \geq \frac{N_1}{N_2}.$$

Let $d \geq 3$. Assume we have $K(\lambda, N_1, N_2) \leq (N_2^{d-3}/\lambda)^{1/2}$ when $\lambda < N_1/N_2^2$. Then

$$K(\lambda, N_1, N_2) \leq N_2^{\epsilon} \frac{N_2^{d-1/(2)}}{N_1^{1/2}} \quad \text{when } \lambda \geq \frac{N_1}{N_2^2}.$$

Note that when $d \geq 3$, Lemmas 3.5 and 3.6 imply Lemma 3.3. In dimension $d = 2$, we use induction (we rely on the so-called parabolic rescaling) to finish the proof of Lemma 3.3.

We end this section with an outline of the structure of the rest of the paper. We show that Proposition 3.1 implies Theorem 1.3 in Section 4. Lemmas 3.2, 3.4, and 3.6 all rely on the exploration of the so-called *transversality*, which essentially allows us to reduce the dimensionality of the problem. We first explore *transversality* in Section 5 and then we prove Lemmas 3.2, 3.4, and 3.6 in Section 6.

The details of the induction procedure (which is nontrivial) that is used to prove Lemma 3.3 in dimension $d = 2$ are given in Section 5. We remark here the proof of Lemma 3.3 relies on Lemma 3.2.

Finally, we prove Lemma 3.5 at the end of Section 7, which, together with Lemma 3.6 will conclude the proof of Lemma 3.3 in dimension $d \geq 3$.

4. Proposition 3.1 implies Theorem 1.3

We first introduce one standard but important tool in the following lemma.

Lemma 4.1 [Bourgain and Demeter 2015; 2017]. *Let $\{g_\alpha\}$ be a family of functions such that* supp $\hat{g}_\alpha$ *are finitely overlapped cubes of length ρ. Let A be bounded convex open set tiled by finitely overlapped cubes Q of side length $\geq \rho^{-1}$. Then for the w_A adapted to A, the following holds*:

$$\fint_A \Big|\sum g_\alpha\Big|^2 w_A \lesssim \sum \frac{1}{m(A)} \int |g_\alpha|^2 w_A.$$

Proof. Since we can sum up the weight function over a finitely overlapping cover $\{Q\}$ of A, that is, $w_A = \sum_{Q \subset A} w_Q$, it suffices to prove the result for $A = Q$. Recall by inequality (1-15), we can cover the whole space $\mathbb{R}^n$ by translations Q' of Q:

$$\begin{aligned}
\fint_Q \Big|\sum g_\alpha\Big|^2 w_Q\, dx &\leq \frac{1}{m(Q)} \sum_{Q'} \int_{Q'} w_Q(Q') \Big|\sum g_\alpha\Big|^2 \\
&\leq \frac{1}{m(Q)} \sum_{Q'} w_Q(Q') \int \Big|\sum g_\alpha\Big|^2 \mu_{Q'}^2 \\
&= \frac{1}{m(Q)} \sum_{Q'} w_Q(Q') \int |\hat{g}_\alpha * \hat{\mu}_{Q'}|^2 \\
&\lesssim \frac{1}{m(Q)} \sum_{Q'} w_Q(Q') \sum_\alpha \int |g_\alpha|^2 \mu_{Q'}^2 \lesssim \frac{1}{m(Q)} \sum_\alpha \int |g_\alpha|^2 w_Q.
\end{aligned}$$

□

Now we can reduce Proposition 3.1 to a bilinear decoupling on two (N_2/N_1)-diameter caps.

Lemma 4.2. *Theorem 1.3 is equivalent to Proposition 3.1.*

Proof. Let f_1, f_2 be as in Theorem 1.3. Then $f_1 = \sum_{|\tau| = N_2/N_1} f_{1,\tau}$ and the $f_{1,\tau}$ are supported on finitely overlapping caps of diameter N_2/N_1.

Since $|f_2|$ is supported in a cap of diameter N_2/N_1, the supports of $\{\widehat{Ef}_{1,\tau} * \widehat{Ef_2}\}_\tau$ are in finitely overlapping cubes of length N_2/N_1. Since the scale of Ω is larger than N_1/N_2, i.e., it contains a ball of radius $> N_1/N_2$, by Lemma 4.1,

$$\fint_\Omega |Ef_1 Ef_2|^2 w_\Omega\, dx \leq \sum_{|\tau| = N_2/N_1} \Big|\fint_\Omega Ef_{1,\tau} Ef_2\Big|^2 w_\Omega\, dx.$$

Now apply Proposition 3.1 for $f_{1,\tau}$ and f_2 for each τ; Theorem 1.3 follows. □

5. Transversality

Let f_1, f_2 be as in Proposition 3.1; then f_1 is supported around $(0,0,\dots,0,1,1)$ and f_2 is supported around $(0,0,\dots,0)$. The main goal of this section is to explore the transversality between $(0,0,\dots,0,1)$ and $(0,0,0,\dots,0)$, or more precisely, the transversality between the unit normal vectors of the truncated parabola at these two points. The main lemma in this section is Lemma 5.1 below, and Corollary 5.7 which essentially follows from Lemma 5.1.

We first introduce some basic notation. Let $\{e_1,\dots,e_d\}$ be the standard basis of $\mathbb{R}^d$. We will encounter caps of radius v around $(0,0,\dots,0)$ and $(0,\dots,0,1,1)$ on the parabola. Note around those two points, when v is small (which is always the case in our work), one may view those caps as their natural projection to $\mathbb{R}^{d-1}$. And their image is essentially a square/cap of radius v. We say that a (v,v^2)-plate is a d-dimensional rectangle with the short side on the e_{d-1}-direction such that its image under the orthogonal projection to R^{d-1} is a $(v\times v\times\cdots\times v\times v^2)$-rectangle.

Lemma 5.1. *Given $|\upsilon|<1$, let f_1 be a function supported on a cap of radius υ, centered at $(0,\dots,0,1,1)$ on the truncated parabola P, and let f_2 be a function supported on a cap of radius υ centered at $(0,\dots,0,0,0)$ on the paraboloid. For a covering $\{\tau_i\}$ of* supp f_i *with (υ,υ^2) plates, with the shorter side on the e_{d-1}-direction, we have the following decoupling inequality: for any $R>\upsilon^{-2}$,*

$$\int |Ef_1Ef_2|^2 w_{B_R} \lesssim \sum_{\tau_1,\tau_2}\int |Ef_{1,\tau_1}Ef_{2,\tau_2}|^2 w_{B_R}. \tag{5-1}$$

Remark 5.2. We thank J. Ramos for pointing out that Lemma 5.1 is a particular case of Proposition 2 in [Ramos 2016]. We still write a proof in this paper for clarity.

Proof. The proof is similar to the proof of the L^4 Strichartz estimate on the one-dimensional torus. From the inequality (1-15), we only need to prove that

$$\int_{B'} |Ef_1Ef_2|^2 \lesssim \sum_{\tau_1,\tau_2}\int |Ef_{1,\tau_1}Ef_{2,\tau_2}|^2\mu_{B'}^2$$

for all translations B' of B_R. Now

$$\int_{B'} |Ef_1Ef_2|^2 \le \sum_{\tau_1,\tau_2,\tau_3,\tau_4}\int_{B'} Ef_{1,\tau_1}Ef_{2,\tau_2}\overline{Ef}_{1,\tau_3}\overline{Ef}_{2,\tau_4}\mu_{B'}^2. \tag{5-2}$$

Let $\xi_i\in\tau_i,\ \xi_i=\big(\xi_{i,1},\dots,\xi_{i,d-1},\sum_{j=1}^{d-1}(\xi_i^j)^2\big)\equiv(\bar{\xi}_i,\xi_{i,d-1},|\bar{\xi}_i|^2+(\xi_i^{d-1})^2),\ i=1,2,3,4.$ We have

$$\begin{aligned} |\bar{\xi}_i| &\lesssim \upsilon, \quad i=1,2,3,4,\\ |\xi_{i,d-1}-1| &\lesssim \upsilon, \quad i=1,3,\\ |\xi_{i,d-1}| &\lesssim \upsilon, \quad i=2,4. \end{aligned} \tag{5-3}$$

Essentially, for any $\tau_1,\tau_2,\tau_3,\tau_4$ such that

$$\int Ef_{1,\tau_1}Ef_{2,\tau_2}\overline{Ef}_{1,\tau_3}\overline{Ef}_{2,\tau_4}\mu_{B'}^2 \neq 0,$$

one must have for some $\xi_i \in \tau_i$,

$$\begin{aligned} \xi_1 - \xi_3 &= \xi_2 - \xi_4 + O(R^{-1}), \\ |\xi_1|^2 - |\xi_3|^2 &= |\xi|_2^2 - |\xi|_4^2 + O(R^{-1}), \end{aligned} \tag{5-4}$$

and the second formula in (5-4) implies

$$(\xi_{1,d-1} - \xi_{3,d-1})(\xi_{1,d-1} + \xi_{3,d-1}) = O(|\xi_2|^2 + |\xi_4|^2) + O(|\bar{\xi}_1|^2 + |\bar{\xi}_3|^2) + O(R^{-1}). \tag{5-5}$$

Plugging into (5-3), one has $|\xi_{1,d-1} - \xi_{3,d-1}| \lesssim \upsilon^2$, which again implies $|\xi_{2,d-1} - \xi_{4,d-1}| \lesssim \upsilon^2$.

To summarize, $\int Ef_{1,\tau_1} Ef_{2,\tau_2} \overline{Ef}_{1,\tau_3} \overline{Ef}_{2,\tau_4} \mu_{B'}^2 \neq 0$ implies the distance between τ_1 and τ_3 and the distance between τ_2 and τ_4 are both bounded by υ^2, which essentially means $\tau_i = \tau_{i+2}$, $i = 1, 2$. Applying this fact to (5-2), Lemma 5.1 follows. □

Remark 5.3. A quantitative version of estimate (5-1) can be stated as follows: assume that the support of f_1 is centered at $(0, 1/K, (1/K)^2)$ rather than $(0, 0, 1)$. From the proof we can attain the same estimate as in (5-1) by introducing an additional constant K,

$$\int |Ef_1 Ef_2|^2 w_{B_R} \lesssim K \sum_{\tau_1,\tau_2} \int |Ef_{1,\tau_1} Ef_{2,\tau_2}|^2 w_{B_R}. \tag{5-6}$$

Indeed, the proof essentially only relies on the fact that for $\xi_i \in \operatorname{supp} f_i$, $i = 1, 2$, the difference between the $d-1$ components is at least $1/K$. Similar arguments also hold for the estimate in Lemma 5.5; see Corollary 5.7 below.

Remark 5.4. We remark that for any $\alpha < \upsilon$, a function which is supported on a cap of radius α can be naturally understood as a function supported on a cap of radius υ.

Lemma 5.1 facilitates the decomposition of caps of radius υ into plates of size (υ, υ^2). We can further decompose those into caps of radius υ^2.

Lemma 5.5. *With the same notation as in Lemma 5.1, $R \geq \upsilon^{-2}$, let $\operatorname{supp} f_i$ be covered by finitely overlapping caps θ_i of radius υ^2, $i = 1, 2$. Then*

$$\int |Ef_1 Ef_2|^2 w_{B_R} \lesssim \upsilon^{-(d-1)} \sum_{|\theta_i| = \upsilon^2} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_R}. \tag{5-7}$$

Proof. Clearly, we need only to prove (5-7) for every ball of radius υ^{-2} contained in B_R, and then sum them together. (This is the same principle of parallel decoupling, Lemma 1.9.)

Fix a pair of (υ, υ^2) plates τ_1, τ_2:

$$\begin{aligned} \int |Ef_{1,\tau_1} Ef_{2,\tau_2}|^2 w_{B_R} &= \int \Bigg| \sum_{\substack{\theta_2 \subset \tau_2 \\ |\theta_2| = \upsilon^2}} Ef_{1,\tau_1} Ef_{2,\theta_2} \Bigg|^2 w_{B_R} \\ &\leq \upsilon^{-(d-1)} \sum_{\substack{\theta_2 \subset \tau_2 \\ |\theta_2| = \upsilon^2}} |Ef_{1,\tau_1} Ef_{2,\theta_2}|^2 w_{B_R} \lesssim \sum_{\substack{\theta_j \subset \tau_j \\ |\theta_j| = \upsilon^2}} |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_R}. \end{aligned} \tag{5-8}$$

The last inequality follows from Lemma 4.1 and Lemma 4.2. □

Remark 5.6. Similar to Remark 5.4, for $\upsilon^2 < \alpha < \upsilon$, a cap of scale υ naturally lies in a cap of scale $\sqrt{\alpha}$. Thus if we let f_1 be a function supported on a cap of radius α centered at $(0, \dots, 0, 1, 1)$ on the paraboloid, and we let f_2 be a function supported on a cap of radius α centered at $(0, \dots, 0, 0, 0)$ on the paraboloid, then by arguing similar to the proof of Lemma 5.5, we have for $R \geq \alpha^{-1}$,

$$\int |Ef_1 Ef_2|^2 w_{B_R} \lesssim \left(\frac{\upsilon}{\alpha}\right)^{d-1} \sum_{|\theta_i|=\alpha} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_R}. \tag{5-9}$$

If we directly use the Hölder inequality for all caps in the support of f_i to estimate as in (5-8), then the interpolation in the proof of Lemma 5.5 will give us a constant υ^{-d} rather than $\upsilon^{-(d-1)}$ in (5-7), since one has υ^{-d} caps for each f_i. The bilinear transversality, i.e., the transversality between $(0, 0, \dots, 0)$ and $(0, \dots, 0, 1, 1)$, helps in reducing the dimension by 1 since in one direction we can use L^4 orthogonality, as shown in Lemma 5.1. Thus here we are able to improve the constant in (5-7) to $\upsilon^{-(d-1)}$.

Corollary 5.7. *Using the same notation as in Lemma 5.1, there exists a constant C such that for any υ, δ, and $R^{-1} \leq \delta \leq \upsilon$,*

$$\int |Ef_1 Ef_2|^2 w_{B_R} \lesssim \left(\frac{\upsilon}{\delta}\right)^{d-1} \left|\frac{\log \delta}{\log \upsilon}\right|^C \sum_{|\theta_i|=\delta} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_R}.$$

Proof. The proof is most clear when $\delta = \upsilon^{2^n}$ for some n. Let us first handle this case and then go to the general case. One may use induction. (This induction, however, does not rely on parabolic rescaling.) If $n = 0$, there is nothing to prove.

Assume the result holds for the case $n = k$. Let us turn to the case $n = k + 1$, where $\delta = \upsilon^{2^{k+1}}$, and so $\delta^{1/2} = \upsilon^{2^k}$; thus by the induction assumption, we have

$$\int |Ef_1 Ef_2|^2 w_{B_R} \lesssim \left(\frac{\upsilon}{\delta^{1/2}}\right)^{d-1} 2^{Ck} \sum_{|\eta_i|=\delta^{1/2}} \int |Ef_{1,\eta_1} Ef_{2,\eta_2}|^2 w_{B_R}. \tag{5-10}$$

Now note that $R \geq (\delta^{-1/2})^2$. By Lemma 5.5, we have for each pair (η_1, η_2) in (5-10) that

$$\int |Ef_{1,\eta_1} Ef_{2,\eta_2}|^2 w_{B_R} \lesssim (\delta^{1/2})^{-(d-1)} \sum_{\substack{\theta_i \subset \eta_i \\ |\theta_i|=\delta}} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_R}. \tag{5-11}$$

The case $n = k + 1$ clearly follows if one plugs (5-11) into (5-10), taking the constant C large enough.

Now we turn to the general case. We only need to work on the case $\upsilon^{2^{n+1}} < \delta < \upsilon^{2^n}$. Recall that previously, when $\delta = \upsilon^{2^n}$, we used induction as $\upsilon \to \upsilon^2 \to \upsilon^{2^2} \to \cdots \to \upsilon^{2^n} = \delta$, and in each step we used Lemma 5.5 to finish the induction $\upsilon^{2^k} \to \upsilon^{2^{k+1}}$.

In the case $\upsilon^{2^{n+1}} < \delta < \upsilon^{2^n}$ we have $\upsilon^{2^n} < \delta^{1/2}$, and we use induction as before for $\upsilon \to \upsilon^2 \to \upsilon^{2^2} \to \cdots \to \upsilon^{2^n}$, and we use (5-9) to use induction again from υ^{2^n} to δ. This ends the proof. ☐

6. Proofs of Lemmas 3.2, 3.4 and 3.6

We are now prepared to use transversality to prove Lemmas 3.2, 3.4, and 3.6. Recall Lemma 3.2 concerns $\widetilde{K}(\lambda, N_1, N_2)$ defined in (3-4). Furthermore, Lemmas 3.4 and 3.6 refer to $K(\lambda, N_1, N_2)$ defined in (3-5).

6A. ***Proof of Lemma 3.2.*** For convenience of notation, we let $\Omega_1 := [0, N_1^2] \times [0, \lambda N_1]^d$. Note that one can use finite overlapped balls of radius N_1^2 to cover Ω_1 since $\lambda \geq N_1$. We want to prove

$$\|Ef_1 Ef_2\|_{L^2_{\mathrm{avg}}(\omega_{\Omega_1})} \lesssim_\epsilon \lambda^{d/2} N_2^\epsilon \frac{N_2^{d-1}}{N_1} \prod_{j=1}^{2} \biggl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{\Omega_1})} \biggr)^{1/2}. \tag{6-1}$$

We first apply Corollary 5.7 with $\delta = N_1^{-2},\ \upsilon = N_2/N_1,\ R = N_1^2$. Note that $\delta \leq \upsilon$. Then we have

$$\begin{aligned}\int |Ef_1 Ef_2|^2 w_{B_{N_1^2}} &\lesssim (N_1N_2)^{d-1} \left| \frac{\log N_1}{\log N_1 - \log N_2} \right|^C \sum_{|\theta_j|=1/N_1^2} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_{N_1^2}} \\ &\lesssim (N_1N_2)^{d-1} N_2^\epsilon \sum_{|\theta_j|=1/N_1^2} \prod_{j=1}^{2} \|Ef_{j,\theta_j}\|^2_{L^4(w_{B_{N_1^2}})}. \end{aligned} \tag{6-2}$$

Remark 6.1. We avoid the case when $N_1 = N_2$, and thus $\ln N_1 - \ln N_2 = 0$, by first decomposing caps of diameter N_2/N_1 into caps of diameter $N_2/(2N_1)$ with loss of a fixed constant, then continuing with the proof as above. In all of what follows, one may assume, without loss of generality, that $N_1 \geq 2N_2$.

Via the principle of parallel decoupling, Lemma 1.9, or by summing different $B_{N_1^2}$ together, we have

$$\int |Ef_1 Ef_2|^2 w_{\Omega_1} \lesssim (N_1N_2)^{d-1} N_2^\epsilon \sum_{|\theta_j|=1/N_1^2} \prod_{j=1}^{2} \|Ef_{j,\theta_j}\|^2_{L^4(w_{\Omega_1})}. \tag{6-3}$$

Next we would like to show that

$$\|Ef_{j,\theta_j}\|^2_{L^4(w_{\Omega_1})} \leq \left(\frac{\lambda}{N_1}\right)^{d/2} \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'|=1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^4(w_{\Omega_1})}. \tag{6-4}$$

It suffices to show

$$\|Ef_{j,\theta_j}\|^2_{L^4_{\mathrm{avg}}(\Omega_1)} \leq \left(\frac{\lambda}{N_1}\right)^{d/2} \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'|=1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^4_{\mathrm{avg}}(w_{\Omega_1})}$$

and sum up as in Lemma 4.1.

Each function $Ef_{j,\theta_j'}$ is Fourier-supported in θ_j', in particular, Fourier-supported in a cylinder of radius $1/(\lambda N_1)$, height $1/N_1^2$, and Ω_1 is tiled by cylinders of radius λN_1, height N_1^2 in the t-direction. The proof of Lemma 4.1 works the same:

$$\|Ef_{j,\theta_j}\|^2_{L^2_{\mathrm{avg}}(\Omega_1)} \lesssim \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'|=1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^2_{\mathrm{avg}}(w_{B_R})} \lesssim \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'|=1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^4_{\mathrm{avg}}(w_{B_R})}.$$

For the L^∞-estimate, we apply the Cauchy–Schwarz inequality:

$$\|Ef_{j,\theta_j}\|^2_{L^\infty(\Omega_1)} \le \left(\frac{\lambda}{N_1}\right)^d \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'| = 1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^\infty(\Omega_1)} \lesssim \left(\frac{\lambda}{N_1}\right)^d \sum_{\substack{\theta_j' \subset \theta_j \\ |\theta_j'| = 1/(\lambda N_1)}} \|Ef_{j,\theta_j'}\|^2_{L^4_{\mathrm{avg}}(w_{\Omega_1})}.$$

The last inequality is an application of Lemma 1.10. Note $f_{\theta_j'}$ is supported in a ball of scale $1/(\lambda N_1)$, and inside a box C of size $(1/N_1^2) \times (1/(\lambda N_1)) \times \cdots \times (1/(\lambda N_1))$. We can make a affine transform of C into a cube Q^* of scale λ_{N_1}, which on the physical side would transform Ω_1 into a cube of scale λN_1. We apply Lemma 1.10 after the affine transformation and then transform back. (Note in that setting, cube is no different than a ball.)

We apply Hölder's inequality to conclude the argument.

6B. ***Proof of Lemma 3.4.*** Let $\lambda \le N_1$. We first note that we can use finitely overlapping balls $B_{\lambda N_1}$ to cover Ω and that $N_2 \lesssim 1$. Applying Corollary 5.7 with $\delta = 1/(\lambda N_1)$ and $\upsilon = N_2/N_1$ we have

$$\begin{aligned}\int |Ef_1 Ef_2|^2 w_{B_{\lambda N_1}} &\lesssim (\lambda N_2)^{d-1} \left|\frac{\log \lambda + \log N_1}{\log N_1 - \log N_2}\right|^C \sum_{|\theta_j| = 1/(\lambda N_1)} \int |Ef_{1,\theta_1} Ef_{2,\theta_2}|^2 w_{B_{\lambda N_1}} \\ &\lesssim (\lambda N_2)^{d-1} N_2^\epsilon \sum_{|\theta_j| = 1/(\lambda N_1)} \prod_{j=1}^2 \|Ef_{j,\theta_j}\|^2_{L^4(w_{B_{\lambda N_1}})}.\end{aligned}$$

With parallel decoupling, Lemma 1.9, the desired estimate follows. (As remarked in Remark 6.1, one can assume $N_1 \ge 2N_2$.)

6C. ***Proof of Lemma 3.6.*** Let $\lambda \le N_1$.

We have the following two cases:

- Case 1: $d = 2$, $N_1 \ge \lambda \ge N_1/N_2$, and $N_2' = (N_1/\lambda)$.
- Case 2: $d \ge 3$, $N_1 \ge \lambda \ge N_1/N_2^2$, and $N_2' = (N_1/\lambda)^{1/2}$.

It is easy to check that we only need to show

$$K(\lambda, N_1, N_2) \lesssim K(\lambda, N_1, N_2') \left(\frac{N_1}{N_2'} \frac{N_2}{N_1}\right)^{(d-1)/2}. \tag{6-5}$$

We claim that

$$\|Ef_1 Ef_2\|_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})} \lesssim \left(\frac{N_1}{N_2'} \frac{N_2}{N_1}\right)^{(d-1)/2} \prod_{j=1}^2 \left(\sum_{|\theta| = N_2'/(N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})}\right)^{1/2}. \tag{6-6}$$

Since $\lambda \le N_1$, we cover $B_{N_1^2}$ with balls of radius λN_1. Thus by parallel decoupling, to prove (6-6), we only need to show

$$\|Ef_1 Ef_2\|_{L^4_{\mathrm{avg}}(w_{B_{\lambda N_1}})} \lesssim \left(\frac{N_1}{N_2'} \frac{N_2}{N_1}\right)^{(d-1)/2} \prod_{j=1}^2 \left(\sum_{|\theta| = 1/(\lambda N_1)} \|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{\lambda N_1}})}\right)^{1/2}. \tag{6-7}$$

Note that since $\lambda N_1 \geq N_1/N_2'$, estimate (6-7) follows from Corollary 5.7 by setting $\delta = N_2/N_1$, $\upsilon = N_2'/N_1$ via interpolation and local constant arguments as in Section 6A.

By the definition of $K(\lambda, N_1, N_2)$, we have that for any θ_1, θ_2 in (6-7),

$$\|Ef_{1,\theta_1} Ef_{2,\theta_2}\|_{L^4_{\mathrm{avg}}(w_{B_{\lambda N_1}})} \lesssim \lambda^{d/2} K(\lambda, N_1, N_2') \prod_{j=1}^{2} \Bigg(\sum_{\substack{|\theta_j'|=1/(\lambda N_1) \\ \theta_j \subset \theta_j}} \|Ef_{j,\theta_j'}\|^2_{L^4_{\mathrm{avg}}(w_\Omega)} \Bigg)^{1/2}. \tag{6-8}$$

Plugging (6-8) into (6-7), clearly (6-5) follows.

7. Induction procedure and proof of Lemma 3.3

To conclude the proof of Proposition 3.1, we are left with the proof of Lemma 3.3. For this lemma the proof relies on induction on N_2. The base case $N_2 \lesssim 1$ is resolved by Lemma 3.4, and by Lemma 3.6 we need only to induct until $\lambda = (N_2)^{d-1}/N_1$.

Let f_1, f_2 be as in Lemma 3.3. Applying Lemma 5.1, taking $v = N_1/N_2$ and $R = N_1^2$, we can decouple the N_2/N_1 caps into $(N_2/N_1, N_2^2/N_1^2)$ plates without any loss; i.e.,

$$\int |Ef_1 Ef_2|^2 w_{B_{N_1^2}} \lesssim \sum_{\tau_1,\tau_2} \int |Ef_{1,\tau_1} Ef_{2,\tau_2}|^2 w_{B^2_{N_1}}. \tag{7-1}$$

Here τ_i are plates as described in Lemma 5.1. We focus on the case when $d = 2$ in $\mathbb{R}^3$; the high-dimensional case will be explained in the end. When $d = 2$, the underlying plates become strips. We start with some preparation before the induction.

7A. *Preliminary preparation for the induction.* We fix a pair of $(N_2/N_1, N_2^2/N_1^2)$ strips τ_1, τ_2 from estimate (7-1). We decompose τ_j into a union of $(N_2/(KN_1)) \times (N_2^2/N_1^2)$ strips $\{s_j\}$.

Using the notation "nonadj" short for nonadjacent, and "adj" short for adjacent, we have

$$\begin{aligned} |Ef_{\tau_j}|^2 &= \sum_{s_j} |Ef_{s_j}|^2 + \sum_{s_j, s_j' \text{adj}} |Ef_{s_j} Ef_{s_j'}| + \sum_{s_j, s_j' \text{nonadj}} |Ef_{s_j} Ef_{s_j'}| \\ &\leq 10 \sum_{s_j} |Ef_{s_j}|^2 + \sum_{s_j, s_j' \text{nonadj}} |Ef_{s_j} Ef_{s_j'}| = I_{j,1} + I_{j,2} \end{aligned} \tag{7-2}$$

and

$$\begin{aligned} \int |Ef_{\tau_1} Ef_{\tau_2}|^2 w_{B_{N_1^2}} &\leq \int \big|(Ef_{\tau_1}^2 - I_{1,1})(Ef_{\tau_2}^2 - I_{2,1})\big| + Ef_{\tau_1}^2 I_{2,1} + Ef_{\tau_2}^2 I_{1,1} + I_{1,1} I_{2,1} w_{B_{N_1^2}} \\ &\lesssim \sum_{s_j, s_j' \text{nonadj}} \int |Ef_{s_1} Ef_{s_1'} Ef_{s_2} Ef_{s_2'}| w_{B_{N_1^2}} + \sum_{s_1, s_2} \int |Ef_{s_1} Ef_{s_2}|^2 w_{B_{N_1^2}}. \end{aligned} \tag{7-3}$$

The last inequality follows from Lemmas 4.1 and 4.2.

The reason why we want to have nonadjacent parts is that we would like transversality (after rescaling) on the other direction. Formula (7-3) will the starting point of our induction.

For the second term in (7-3), we will later directly use induction (not relying on parallel rescaling) on N_2 and reduce everything to the known base case $N_2 = 1$.

For the first term, using Cauchy–Schwarz

$$\int |Ef_{s_1} Ef_{s_1'} Ef_{s_2} Ef_{s_2'}| w_{B_{N_1^2}} \le \left(\int |Ef_{s_1} Ef_{s_1'}|^2 w_{B_{N_1^2}} \right)^{1/2} \left(\int |Ef_{s_2} Ef_{s_2'}|^2 w_{B_{N_1^2}} \right)^{1/2}. \tag{7-4}$$

We point out here that in what follows we do not rely on the bilinear transversality between s_1 and s_2 (or s_1 and s_2'), which is already handled in Lemma 5.1. Instead we will rely on the bilinear transversality between s_1 and s_1' (or s_2, s_2'), since they are not adjacent. This transversality is most clear when one applies parabolic rescaling.

Let us now turn to the term $\int |Ef_{s_2} Ef_{s_2'}|^2 w_\Omega$, when s_2, s_2' are nonadjacent. The term with s_1, s_1' is handled similarly, though one may need to rotate the coordinates.

Finally we point out here that K will be chosen large later and any (fixed) power of K will not impact the final estimate. In particular, in the following estimates we will not worry about losing powers of K.

Without loss of generality, we assume

- s_2 is the strip $\{(a_1, a_2, a_1^2 + a_2^2) : |a_1| \le N_2^2/N_1^2,\ |a_2| \le N_2/KN_1\}$,
- s_2' is the strip $\{(b_1, b_2, b_1^2 + b_2^2) : |b_1| \le N_2^2/N_1^2,\ |b_2 - CN_2/KN_1| \le N_2/KN_1\},\ C \ge 10$. (Here 10 is of course just some universal constant.)

7B. ***Parabolic rescaling.*** The next step, parabolic scaling, is standard in decoupling-type results; we give the details here for the convenience of the reader.

Note s_2, s_2' lie on the same N_2/N_1 cap. We rescale the N_2/N_1 cap to radius 1. By a slight abuse of notation, we regard f_{s_i} as a function depending only on two variables $(\xi_{i,1}, \xi_{i,2})$. For convenience of notation, we let $h_1 = f_{s_2},\ h_2 = f_{s_2'}$. Let also $g_i(\eta_{i,1}, \eta_{i,2}) := h_i\big((N_2/N_1)\eta_{i,1}, (N_2/N_1)\eta_{i,2}\big)$.

Now

- g_1 is supported in the strip $\{(a_1, a_2, a_1^2 + a_2^2) : |a_1| \le N_2/N_1,\ |a_2| \le 1/K\}$,
- g_2 is supported in the strip $\{(b_1, b_2, b_1^2 + b_2^2) : |b_1| \le N_2/N_1,\ |b_2 - C/K| \le 1/K\},\ C \ge 10$.

Note g_1, g_2 are supported on a pair of transverse $(N_2/N_1) \times 1$ strips[1] due to the nonadjacency of s_2, s_2'. We point out here the transversality between g_1, g_2 is not as in the assumption of Lemma 5.1, but it is in the sense of Remark 5.3, which usually causes a loss of K in the estimate, but this does not matter.

The *parabolic scaling* says the following:

Claim 7.1. *Let*

$$Eg_i(y_1, y_2, y_3) = Eh_i((N_1/N_2)y_1, (N_1/N_2)y_2, (N_1^2/N_2^2)y_3),$$

let D be domain in $\mathbb{R}^3$ and let

$$\widetilde{D} := \{(y_1, y_2, y_3) : (N_1/N_2)y_1, (N_1/N_2)y_2, (N_1^2/N_2^2)y_3 \in D\}.$$

[1] Strictly speaking, we need them to support on a pair of $(N_2/N_1) \times (1/100)$ strips; we neglect this technical point here.

Then it follows from a standard change of variables technique that the following two estimates, with the same constant A, are equivalent:

$$\|Eh_1 Eh_2\|_{L^2_{\mathrm{avg}}(w_D)} \lesssim A \prod_{j=1}^{2} \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{s_j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_D)} \Bigr)^{1/2}, \tag{7-5}$$

$$\|Eg_1 Eg_2\|_{L^2_{\mathrm{avg}}(w_{\widetilde{D}})} \lesssim A \prod_{j=1}^{2} \Bigl(\sum_{|\tilde{\theta}|=1/(\lambda N_2)} \|Eg_{j,\tilde{\theta}}\|^2_{L^4_{\mathrm{avg}}(w_{\widetilde{D}})} \Bigr)^{1/2}. \tag{7-6}$$

We then concentrate on (7-6).

Take $D = B_{N_1^2}$; then $\widetilde{D} = [0, N_2^2] \times [0, N_1 N_2]^2$. (Here, without loss of generality, we regard $B_{N_1^2}$ as $[0, N_1^2]^3$.) For convenience of notation, we set $\widetilde{\Omega} = [0, N_2^2] \times [0, N_1 N_2]^2$. The parabolic rescaling gives:

Lemma 7.2. *Assume g_1, g_2 are two general functions defined on the parabola. Let g_1 be supported in a strip of size $(N_2/N_1) \times 1$ around $(0,0,0)$, and g_2 be supported in a strip of size $(N_2/N_1) \times 1$ around $(0,1,1)$. If for some constant A, one has* (*for all such g_1, g_2*)

$$\|Eg_1 Eg_2\|_{L^2_{\mathrm{avg}}(w_{\widetilde{\Omega}})} \lesssim A \Bigl(\sum_{|\tilde{\theta}|=1/(\lambda N_2)} \|Eg_{j,\tilde{\theta}}\|^2_{L^4_{\mathrm{avg}}(w_{\widehat{\Omega}})} \Bigr)^{1/2}, \tag{7-7}$$

then for the same constant A, one has

$$\|Ef_{s_2} Ef_{s_2'}\|_{L^2_{\mathrm{avg}}(w_{B_{N_1^2}})} \lesssim K^C A \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{s_2,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})} \Bigr)^{1/2} \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{s_2',\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})} \Bigr)^{1/2}. \tag{7-8}$$

Remark 7.3. After rescaling, the relevant g_1, g_2 should be supported around $(0,0,0)$ and $(0, 1/K, 1/K^2)$ rather than $(0,0,0)$ and $(0,0,1)$. We state our lemma for g_1, g_2 supported around $(0,0,0)$ and $(0,1,1)$ to be consistent with the statement in Lemma 5.1. This causes a loss of K^C, but we emphasize again that any loss due to a power of K would be irrelevant in the proof.

We end this section by introducing some notation.

Let g_1, g_2 be as in Lemma 7.2. We define $A(\lambda, N_1, N_2)$ to be the best constant such that

$$\|Eg_1 Eg_2\|_{L^2_{\mathrm{avg}}(w_{\widetilde{\Omega}})} \lesssim A(\lambda, N_1, N_2) \Bigl(\sum_{|\tilde{\theta}|=1/(\lambda N_2)} \|Eg_{j,\tilde{\theta}}\|^2_{L^4_{\mathrm{avg}}(w_{\widehat{\Omega}})} \Bigr)^{1/2}. \tag{7-9}$$

Then we can restate Lemma 7.2.

Lemma 7.4. *For $j = 1, 2$, we have*

$$\|Ef_{s_j} Ef_{s_j'}\|_{L^2_{\mathrm{avg}}(w_{B_{N_1^2}})} \lesssim K^C A(\lambda, N_1, N_2) \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{s_j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})} \Bigr)^{1/2} \Bigl(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{s_j',\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})} \Bigr)^{1/2}. \tag{7-10}$$

7C. ***The induction procedure.***

7C1. *Before induction.* Now we are ready to start the induction for the proof of Lemma 3.3. We emphasize here the induction is on N_2 (though mixed with induction on K). Note we are now in dimension $d=2$.

We need to show that for all $1 \le N_2 \le N_1$ and $\lambda \le N_1$, one has

$$K(\lambda, N_1, N_2) \lesssim N_2^{\epsilon}\left(\frac{1}{\lambda} + \frac{N_2}{N_1}\right)^{1/2}.$$

Note the base case $N_2 = 1$ is already established in Lemma 3.4. And with Lemma 3.6, we need only to perform induction until $\lambda = N_2/N_1$.

We will work on $A(\lambda, N_1, N_2)$ defined in (7-9) to explore the transversality between nonadjacent strips. The induction process is two-fold in some sense. We will induct on N_2 to better understand $K(\lambda, N_1, N_2)$. In turn we find more information about $A(\lambda, N_1, N_2)$, which gives a better understanding of $K(\lambda, N_1, N_2)$.

This is a final summary before we start the induction. Recall, we have (7-1) and (7-3); thus we have

$$\int |Ef_1 Ef_2|^2 w_{B_{N_1^2}} \lesssim \int \sum_{s_j, s_j' \text{ nonadj}} \int |Ef_{s_1} Ef_{s_1'} Ef_{s_2} Ef_{s_2'}| w_{B_{N_1^2}} + \int_{s_1, s_2} \int |Ef_{s_1} Ef_{s_2}| w_{B_{N_1^2}}. \tag{7-11}$$

Also recall that s_1, s_1', s_2, s_2' are all $(N_2/N_1)^2 \times (N_2/KN_1)$ strips. The second term can be easily handled by direct induction, (which is not the main point of the induction procedure explained later). Indeed, if there were only the second term in (7-11), since s_1, s_2 are both contained in caps of radius (N_2/KN_1), then (7-11) already reduces the decoupling problem for f_i supported in caps of size N_2/N_1 into the decoupling problem for f_i supported in caps of size N_2/KN_1, which reduce N_2 to N_2/K.

We will focus on the first term of (7-11). The Hölder inequality gives

$$\int |Ef_{s_1} Ef_{s_1'} Ef_{s_2} Ef_{s_2'}| w_{B_{N_1^2}} \le \prod_{j=1}^{2} \left(\int |Ef_{s_j} Ef_{s_j'}|^2 w_{B_{N_1^2}} \right)^{1/2}. \tag{7-12}$$

Estimate (7-12) is the starting point of the analysis in the following subsections.

We summarize in the lemma below how (7-12) and (7-11) come together to highlight the relevance of $A(N_1, N_2, \lambda)$ in the induction procedure.

Lemma 7.5. *When $\lambda \le N_1/N_2$ and $\lambda \le N_1$, we have*

$$K(N_1, N_2, \lambda) \lesssim K^C \frac{1}{\lambda} A(N_1, N_2, \lambda) + K(N_1, N_2/K, \lambda). \tag{7-13}$$

Note that the assumption of Lemma 7.5 always holds during the induction procedure to prove Lemma 3.3.

Proof of Lemma 7.5. Applying Lemma 7.4, we have

$$\|Ef_{s_j} Ef_{s_j'}\|_{L^2_{\text{avg}}(w_{B_{N_1^2}})} \lesssim K^C A(N_1, N_2, \lambda) \left(\sum_{|\theta| = 1/(\lambda N_1)} \|Ef_{s_j,\theta}\|^2_{L^4_{\text{avg}}(w_{B_{N_1^2}})} \right)^{1/2} \left(\sum_{|\theta| = 1/(\lambda N_1)} \|Ef_{s_j',\theta}\|^2_{L^4_{\text{avg}}(w_{B_{N_1^2}})} \right)^{1/2}. \tag{7-14}$$

Plugging (7-14) into (7-12), and then plugging into (7-11), we derive

$$\|Ef_1 Ef_2\|_{l^2(w_{B_{N_1^2}})} \lesssim K^C \lambda \left(\frac{1}{\lambda}\right)^{1/2} \prod_{i=1}^{2} \left(\sum_{|\theta|=1/(\lambda N_1)} \|Ef_{i,\theta}\|^2_{L^4_{\text{avg}}(w_{B_{N_1^2}})} \right)^{1/2} + \left(\sum_{|\theta|=N_2/(\lambda K N_1)} \|Ef_{i,\theta}\|^2_{L^4_{\text{avg}}(w_{B_{N_1^2}})} \right)^{1/2}. \tag{7-15}$$

Thus we derive

$$\lambda K(N_1, N_2, \lambda) \lesssim K^C A(N_1, N_2, \lambda) + \lambda K(N_1, N_2/K, \lambda). \tag{7-16}$$

Therefore, Lemma 7.5 follows. □

Now we are ready to start with the induction procedure on N_2. We emphasize again that by Lemma 3.6 we only need to consider the case $\lambda \le N_1/N_2$.

7C2. *First induction: Case $N_2^2 \le N_1$.* It will become clear in the following proof why we choose the first splitting point at $N_1 = N_2^2$. We start with an estimate for $A(\lambda, N_1, N_2)$. We have:

Lemma 7.6. *When $N_2 \le N_1^2$, $\lambda \le N_1$, $\lambda \le N_1/N_2$,*

$$A(\lambda, N_1, N_2) \lesssim \lambda^{1/2} \equiv \lambda\lambda^{-1/2}. \tag{7-17}$$

Assuming Lemma 7.6 for the moment, let us finish the proof of Lemma 3.3 when $N_1 \ge N_2^2$. Applying Lemma 7.6 with Lemma 7.5, we derive

$$K(N_1, N_2, \lambda) \lesssim K^C \lambda \left(\frac{1}{\lambda}\right)^{1/2} + K(N_1, N_2/K, \lambda) \tag{7-18}$$

when $N_1 \ge N_2^2$ and $\lambda \le N_1/N_2$. Choosing $1 \ll K \sim N_2^{\epsilon^{10}}$, performing induction on N_2 again, and recalling that the case $N_2 \lesssim 1$ is covered by Lemma 3.4, we have Lemma 3.3 follows when $N_1 \ge N_2^2$.

Now, we turn to the proof of Lemma 7.6.

Proof of Lemma 7.6. Since $N_1 \le N_2^2$, we have $N_2/N_1 \le 1/N_2$. (It is exactly because of this that we decided our first splitting point $N_1 \le N_2^2$). Thus, the support of g_1, g_2 appearing in (7-9) is (contained in) strips of size $(1/N_2) \times 1$. Thus, in a ball of radius N_2^2, we have

$$\int |Eg_1 Eg_2| w_{B_{N_2^2}} \lesssim \sum_{\substack{|\theta_i|=1/N_2 \\ \theta_i \subset \operatorname{supp} g_i}} \int |E_{g_{1,\theta_1}} Eg_{2,\theta_2}| w_{B_{N_2^2}}. \tag{7-19}$$

The proof of (7-19) is essentially the same as the proof of Lemma 5.1 and we leave it to reader.

Note one can use balls $B_{N_2^2}$ to cover $\widetilde{\Omega} := [0, N_2^2] \times [0, N_1 N_2]^2$ (since $N_1 \ge N_2$), thus we extend (7-19) to

$$\int |Eg_1 Eg_2| w_{\widetilde{\Omega}} \lesssim \sum_{\substack{|\theta_i|=1/N_2 \\ \theta_i \subset \operatorname{supp} g_i}} \int |E_{g_{1,\theta_1}} Eg_{2,\theta_2}| w_{\widetilde{\Omega}}. \tag{7-20}$$

We claim for any fixed θ_1, θ_2, one has

$$\|Eg_{1,\theta_1}Eg_{2,\theta_2}\|_{L^2(w_{\widetilde{\Omega}})} \lesssim \lambda\lambda^{-1/2}\prod_{i=1}^{2}\Bigg(\sum_{\substack{\tilde{\theta}_i\subset\theta_i\\ \|\tilde{\theta}_i=1/(\lambda N_2)\|}}\|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{\widetilde{\Omega}})}\Bigg)^{1/2}. \tag{7-21}$$

Plugging (7-21) into (7-20), we have

$$A(N_1,N_2,\lambda) \lesssim \lambda\left(\frac{1}{\lambda}\right)^{1/2}, \tag{7-22}$$

and then Lemma 7.6 follows.

Now we are left with the proof of (7-21). Let $N_1' = N_2,\ N_2' = N_2^2/N_1 \lesssim 1$. When $N_1' = N_2 \le \lambda$, recall the definition of $\widetilde{K}(\lambda, N_1, N_2)$ in (3-4) and apply Lemma 3.2. Then we have

$$K(N_1',N_2',\lambda) \lesssim (N_2')^{\epsilon}\left(\frac{N_2'}{N_1'}\right)^{1/2} \lesssim \left(\frac{1}{\lambda}+\frac{N_2}{N_1}\right)^{1/2} \lesssim \lambda^{-1/2}. \tag{7-23}$$

The last inequality in (7-23) follows because we always have $\lambda \le N_1/N_2$ in the whole induction process. Note (7-23) implies

$$\begin{aligned}\|Eg_{1,\theta_1}Eg_{2,\theta_2}\|_{L^2(w_{[0,N_2^2]\times[0,\lambda N_2]^2})} \\ \lesssim \lambda\widetilde{K}(N_1',N_2',\lambda)\prod_{i=1}^{2}\Bigg(\sum_{\substack{\tilde{\theta}_i\subset\theta_i\\ \|\tilde{\theta}_i=1/(\lambda N_2)\|}}\|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{[0,N_2^2]\times[0,\lambda N_2]^2})}\Bigg)^{1/2}.\end{aligned} \tag{7-24}$$

Since $\lambda \le N_1$, (which is also always the case during the induction process), $\widetilde{\Omega}$ can be covered by the translations of $[0, N_2^2]\times[0, \lambda N_2]$; thus (7-24) implies (7-21) by parallel decoupling, Lemma 1.9.

When $\lambda \le N_1'$, since $N_2' \lesssim 1$, by Lemma 3.4, we have

$$K(\lambda, N_1', N_2') \lesssim \lambda^{-1/2}. \tag{7-25}$$

Thus,

$$\|Eg_{1,\theta_1}Eg_{2,\theta_2}\|_{L^2(w_{B_{N_2^2}})} \lesssim \lambda\lambda^{-1/2}\prod_{i=1}^{2}\Bigg(\sum_{\substack{\tilde{\theta}_i\subset\theta_i\\ \|\tilde{\theta}_i=1/(\lambda N_2)\|}}\|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{B_{N_2^2}})}\Bigg)^{1/2}. \tag{7-26}$$

Since one can use $B_{N_2^2}$ and its translations to cover $\widetilde{\Omega}$, (7-26) implies (7-21) by parallel decoupling, Lemma 1.9. $\square$

7C3. *Second induction: Case $N_2^{3/2} \le N_1 \le N_2^2$.*

Lemma 7.7. *When $N_2^{3/2} \le N_1 \le N_2^2$, $\lambda \le N_1$ and $\lambda \le N_1/N_2$, we have*

$$A(\lambda, N_1, N_2) \lesssim \lambda^{1/2} \equiv \lambda\lambda^{-1/2}. \tag{7-27}$$

Clearly, using Lemma 7.5 and arguing as in Section 7C2, Lemma 3.3 follows from Lemma 7.7 when $N_2^{3/2} \le N_1 \le N_2^2$.

Now we are left with proof of Lemma 7.7, i.e., the estimate (7-27). We will prove that estimate (7-27), in the case $N_2^{3/2} \le N_1 \le N_2^2$, follows from the fact that Lemma 3.3 holds when $N_2^2 \ge N_1$ (given Lemma 3.2).

Proof of Lemma 7.7. The proof starts similarly to the proof of Lemma 7.6; note now we have $N_2/N_1 \ge 1/N_2$. As we derived in (7-19), we have in a ball of radius N_1^2/N_2^2,

$$\int |Eg_1 Eg_2| w_{B_{(N_1/N_2)^2}} \lesssim \sum_{\substack{|\theta_i|=N_2/N_1 \\ \theta_i \subset \operatorname{supp} g_i}} \int |E_{g_1,\theta_1} Eg_{2,\theta_2}| w_{B_{(N_1/N_2)^2}}. \tag{7-28}$$

Note one can use $B_{(N_1/N_2)^2}$ and its translations to cover $\widetilde{\Omega}$; thus we have

$$\int |Eg_1 Eg_2| w_{\widetilde{\Omega}} \lesssim \sum_{\substack{|\theta_i|=N_2/N_1 \\ \theta_i \subset \operatorname{supp} g_i}} \int |E_{g_1,\theta_1} Eg_{2,\theta_2}| w_{\widetilde{\Omega}}. \tag{7-29}$$

The following procedure is essentially the same as in the first induction. Note that to prove (7-27) we only need to further show that for fixed θ_1, θ_2,

$$\|Eg_{1,\theta_1} Eg_{2,\theta_2}\|_{L^2(w_{\widetilde{\Omega}})} \lesssim \lambda\lambda^{-1/2} \prod_{i=1}^{2} \Biggl(\sum_{\substack{\tilde{\theta}_i \subset \theta_i \\ |\tilde{\theta}_i|=1/(\lambda N_2)}} \|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{\widetilde{\Omega}})} \Biggr)^{1/2}, \tag{7-30}$$

where now $|\theta_i| = N_2/N_1$.

Let $N_1' = N_2$ and $N_2' = N_2^2/N_1$; note we have $N_1' \ge (N_2')^2$ since $N_1 \ge N_2^{3/2}$. When $\lambda \ge N_1'$, we have by Lemma 3.2

$$\|Eg_{1,\theta_1} Eg_{2,\theta_2}\|_{L^2(w_{[0,N_2^2]\times[0,\lambda N_2]^2})} \lesssim \lambda \biggl(\frac{N_2'}{N_1'}\biggr)^{-1/2} \prod_{i=1}^{2} \Biggl(\sum_{\substack{\tilde{\theta}_i \subset \theta_i \\ |\tilde{\theta}_i|=1/(\lambda N_2)}} \|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{[0,N_2]\times[0,\lambda N_2]^2})} \Biggr)^{1/2}. \tag{7-31}$$

Since one can use $[0, N_2^2] \times [0, \lambda N_2]^2$ to cover $\widetilde{\Omega}$, (7-30) follows from (7-31) (note that $N_2'/N_1' = N_2/N_1 \le \lambda^{-1}$).

When $\lambda \le N_1'$, since one can use $B_{N_2^2}$ to cover $\widetilde{\Omega}$, to prove (7-30), we need only show

$$\|Eg_{1,\theta_1} Eg_{2,\theta_2}\|_{L^2(w_{B_{N_2^2}})} \lesssim \lambda\lambda^{-1/2} \prod_{i=1}^{2} \Biggl(\sum_{\substack{\tilde{\theta}_i \subset \theta_i \\ \|\tilde{\theta}_i=1/(\lambda N_2)\|}} \|Eg_{i,\tilde{\theta}_i}\|_{L^4(w_{B_{N_2^2}})} \Biggr)^{1/2}, \tag{7-32}$$

which is equivalent to $K(N_1', N_2', \lambda) \le 1/\lambda$. But recall that $N_1' \ge (N_2')^2$, thus this is exactly what we proved in first induction; i.e., Lemma 3.3 holds when $N_1 \ge N_2^2$. $\square$

7C4. *Later inductions and the conclusion of the induction process.* Recall that the *first induction* covers the case $N_1 \ge N_2^2$ and the *second inductions* covers the case $N_2^\alpha \le N_1 \le N_2^2$, $\alpha = \frac{3}{2}$. The goal now is to use induction to cover the case $N_2^\alpha \le N_1$, all the way to $\alpha = 1$. The arguments here are similar to those for the *second induction* presented in Section 7C3. Let $N_1' = N_2$, $N_2' = N_2^2/N_1$; then $N_1' \ge (N_2')^\alpha$ is equivalent to $N_1 \ge N_2^{(2\alpha-1)/\alpha}$. Once we show that Lemma 3.3 holds when $N_2^\alpha \le N_1 \le N_2^2$, we will be able to extend Lemma 7.7 to the case when $N_2^{(2\alpha-1)/\alpha} \le N_1$, which in turn proves that Lemma 3.3 holds when $N_2^{(2\alpha-1)/\alpha} \le N_1 \le N_2^2$. The induction would not end until $\alpha = 1$. We finally point out that only an induction with finite steps is involved.

To show Lemma 3.3 for a fixed ϵ_0, we may pick an $\tilde{\epsilon} \ll \epsilon_0$, and then perform the induction for $\tilde{\epsilon}$ as above.

After we prove Lemma 3.3 for $N_1 \ge N_2^{1+\tilde{\epsilon}}$, we are left with the case $N_1 \le N_2^{1+\tilde{\epsilon}}$. We first use the Hölder inequality to shrink the size of the cap from N_2/N_1 to $N_2^{1-2\tilde{\epsilon}}/N_1$, which only gives a loss of $N_2^{C\tilde{\epsilon}} \ll N_2^{\epsilon_0}$. Then we use Lemma 3.3 in the case $N_1 \ge N_2^{1+\tilde{\epsilon}}$ again.

Thus, Lemma 3.3 holds for all the cases for our fixed ϵ_0.

7D. ***The high-dimension case.*** To handle the case $d \ge 3$, we are left with the proof of Lemma 3.5. The proof is indeed similar to previous arguments in this section and easier. The proof relies on the linear decoupling estimate in [Bourgain and Demeter 2015].

As mentioned earlier, applying Lemma 5.1, taking $v = N_2/N_1$ and $R = N_1^2$, we can decouple the N_2/N_1 caps into $(N_2/N_1, N_2^2/N_1^2)$ plates without any loss, i.e., (7-1). However, since we are in the case $\lambda \le N_1/N_2^2$, indeed $N_2^2/N_1^2 \lesssim 1/(\lambda N_1)$, we only need a weaker version of (7-1); i.e., we only want to decouple the N_2/N_1 caps into $(N_2/N_1, 1/(\lambda N_1))$ plates:

$$\int |Ef_1 Ef_2|^2 w_{B_{N_1^2}} \lesssim \sum_{\tau_1,\tau_2} \int |Ef_{1,\tau_1} Ef_{2,\tau_2}|^2 w_{B^2_{N_1}}. \tag{7-33}$$

Here τ_i are $(N_2/N_1, 1/(\lambda N_1))$ plates as described in Lemma 5.1. Note (7-33) follows from (7-1).

Now, for each τ_i fixed, we further decouple τ_i into $(1/N_1, 1/(\lambda N_1))$ plates via linear decoupling in [Bourgain and Demeter 2015], here recalled in (1-11). Note direct application of linear decoupling in dimension d gives us

$$\|Ef_{\tau_i}\|_{L^4(w_{B_{N_1^2}})} \lesssim N_2^{\epsilon} (N_2^2)^{d/4-(d+2)/8} \Bigg(\sum_{v_i \subset \tau_i} \|Ef_{v_i}\|^2_{L^4(w_{B_{N_1^2}})} \Bigg)^{1/2}. \tag{7-34}$$

However, we are able to use (1-11) when the dimension is $d-1$ rather than d, because our plates are so thin (of scale $1/(\lambda N_1) \le 1/N_1$), which reduces the dimension by 1. Indeed, linear decoupling (1-11) not only works for those functions which are exactly supported in parabola P but also those which are supported in an N_1^{-2} neighborhood of P. This is consistent with the uncertainty principle, since in physical space we are of scale N_1^2, and in frequency space any scale of N_1^{-2} cannot be differentiated. Since our plates are so thin, of scale $1/(\lambda N_1) \le N_1^{-2}$, one could indeed view them as N_1^{-2} neighborhoods of some $(d-1)$-dimensional parabola. To be more specific, use τ_2 as example, since τ_2 is supported at the origin. Let $\pi_t^{-1}(\tau_2)$ be the pull back image of τ_2 to the paraboloid. The Fourier inverse transform of Ef_{τ_2} is supported on $\pi_t^{-1}(\tau_2)$. One can see that if we project along the x_1-axis, the projection image of $\pi_t^{-1}(\tau_2)$ is the $(1/(\lambda N_1))^2$-neighborhood of a $(d-1)$-dimensional paraboloid (a piece of length N_2/N_1).

Now, applying $(d-1)$-dimensional linear decoupling, we improve (7-34) into

$$\|Ef_{\tau_i}\|_{L^4(w_{B_{N_1^2}})} \lesssim N_2^{\epsilon} (N_2^2)^{(d-1)/4-(d+1)/8} \Bigg(\sum_{v_i \subset \tau_i} \|Ef_{v_i}\|^2_{L^4(w_{B_{N_1^2}})} \Bigg)^{1/2}, \tag{7-35}$$

where v_i are $(1/N_1, 1/(\lambda N_1))$ plates.

Finally, similarly to the derivation of (6-4), we decouple v_i into caps of radius $1/(\lambda N_1)$,

$$\|Ef_{v_i}\|^4_{L^4(w_{B_{N_1^2}})} \lesssim \lambda^{(d-1)}\biggl(\sum_{\theta_i\subset v_i}\|Ef_{\theta_i}\|^2_{L^4(w_{B_{N_1^2}})}\biggr)^2. \tag{7-36}$$

We remark that each v_i can be covered by λ^{d-1} rather than λ^d caps of radius $1/(\lambda N_1)$. Plugging (7-36) into (7-35), then plugging it into (7-33), we derive

$$\|Ef_1 Ef_2\|_{L^2_{\mathrm{avg}}(w_{B_{N_1^2}})} \le \lambda^{d-1/2}N_2^{(d-3)/2}\prod_{j=1}^{2}\biggl(\sum_{|\theta|=1/(\lambda N_1)}\|Ef_{j,\theta}\|^2_{L^4_{\mathrm{avg}}(w_{B_{N_1^2}})}\biggr)^{1/2}. \tag{7-37}$$

Thus, the desired estimate for $K(\lambda, N_1, N_2)$ follows.

Appendix: Sharpness of Theorems 1.1 and 1.3

The sharpness (up to N_2^{ϵ}) of Theorem 1.3 is provided by the following examples. One can also rescale those examples to show the sharpness of Theorem 1.1.

We take

$$Ef_1 = \sum_{\substack{\xi\in\Lambda_{\lambda N_1}\\ |\xi|\le N_2/N_1}} e^{2\pi i(\xi\cdot x+|\xi|^2 t)}$$

and $f_2 = f_1(\cdot - (1, 0, \dots, 0))$. Then $|Ef_1|$ is about $(\lambda N_2)^d$ at $B(0, N_1/N_2)$ in $\mathbb{R}^{d+1}$. Note that it follows from the uncertainty principle that it is locally constant in any ball of size N_1/N_2 and one can easily compute $|Ef_1(0)| \sim (\lambda N_2)^d$. Also note $|Ef_1|$ has periodicity around λN_1 in all components of x (not necessarily in t). The same is true for $|Ef_2|$. Thus,

$$\|Ef_1 Ef_2\|^2_{L^2(w_\Omega)} \gtrsim (\lambda N_2)^{4d}\left|B\left(0, \frac{N_1}{N_2}\right)\right|(\lambda N_1)^d \gtrsim \lambda^{5d}N_1^{2d+1}N_2^{3d-1}.$$

Each cap θ_j of radius $1/(\lambda N_1)$ contains at most one point $\xi \in \Lambda_{\lambda N_1}$. Hence $\|Ef_{j,\theta_j}\|^4_{L^4(w_\Omega)} \lesssim |\Omega| = N_1^2(\lambda N_1)^{2d}$:

$$\Pi_{j=1}^2\biggl(\sum_{|\theta_j|=1/(\lambda N_1)}\|Ef_{j,\theta_j}\|^2_{L^4(w_\Omega)}\biggr) \lesssim (\lambda N_2)^{2d}N_1^2(\lambda N_1)^{2d} \lesssim \lambda^{4d}N_1^{2d+2}N_2^{2d}.$$

This example shows that the term with N_2^{d-1}/N_1 is sharp for both $d = 2$ and $d \ge 3$.

When $d = 2$, we consider the example when

$$Ef_1 = \sum_{\substack{\xi\in\Lambda_{\lambda N_1}\\ \xi_1=1,\,|\xi_2|\le 1/N_1}} e^{2\pi i(\xi\cdot x+|\xi|^2 t)}, \quad Ef_2 = \sum_{\substack{\xi\in\Lambda_{\lambda N_1}\\ \xi_1=0\,|\xi_2|\le 1/N_1}} e^{2\pi i(\xi\cdot x+|\xi|^2 t)}.$$

$|Ef_1|$ is about λ in the box of height N_1^2 (i.e., the t-direction), width N_1 (i.e., the-x_2 direction) and length $(\lambda N_1)^2$ (i.e., the x_1-direction) centered at origin. $|Ef_2|$ is the same size in the same box. Moreover,

Ef_1 and Ef_2 both have periodicity around λN_1 in x_2:

$$\|Ef_1 Ef_2\|^2_{L^2(w_\Omega)} \gtrsim \lambda^4 N_1^2 \cdot N_1 \cdot (\lambda N_1)^2 \cdot \lambda N_1 \gtrsim \lambda^7 N_1^6.$$

As calculated previously, $\|Ef_{j,\theta_j}\|^4_{L^4(w_\Omega)} = |\Omega|$, so

$$\Pi_{j=1}^2 \Bigg(\sum_{|\theta_j|=1/(\lambda N_1)} \|Ef_{j,\theta_j}\|^2_{L^4(w_\Omega)} \Bigg) \lesssim \lambda^2 \cdot |\Omega| \lesssim \lambda^6 N_1^6.$$

This example shows that when $d = 2$, the term with $1/\lambda$ is sharp.

When $d \geq 3$, we consider the example with

$$Ef_1 = \sum_{\substack{\xi \in \Lambda_{\lambda N_1}, \xi_1 = 1 \\ |(\xi_2, \dots, \xi_d)| \leq N_2/N_1}} e^{2\pi i (\xi \cdot x + |\xi|^2 t)}, \quad Ef_2 = \sum_{\substack{\xi \in \Lambda_{\lambda N_1}, \xi_1 = 0 \\ |(\xi_2, \dots, \xi_d)| \leq N_2/N_1}} e^{2\pi i (\xi \cdot x + |\xi|^2 t)}.$$

Notice that we construct the example in $d \geq 3$ differently; the support of f_j is in a thin plate of radius N_2/N_1 instead of $1/N_1$, as in two-dimensional example.

$|Ef_1|$ is about $(\lambda N_2)^{d-1}$ in a box of size $(N_1/N_2) \times \cdots \times (N_1/N_2) \times (N_1/N_2)^2 \times (\lambda N_1)^2$. $|Ef_2|$ is about $(\lambda N_2)^{d-1}$ in the same box. Both Ef_1 and Ef_2 have periodicity around λN_1 in the x_2-, …, x_d-directions:

$$\|Ef_1 Ef_2\|^2_{L^2(w_\Omega)} \gtrsim (\lambda N_1)^{4(d-1)} \left(\frac{N_1}{N_2}\right)^{d+1} (\lambda N_1)^2 (\lambda N_1)^{d-1} \gtrsim \lambda^{5d-3} N_1^{2d+2} N_2^{3d-5}$$

and

$$\Pi_{j=1}^2 \Bigg(\sum_{|\theta_j|=1/(\lambda N_1)} \|Ef_{j,\theta_j}\|^2_{L^4(w_\Omega)} \Bigg) \lesssim (\lambda N_2)^{2(d-1)} \cdot |\Omega| \lesssim \lambda^{4d-2} N_1^{2d+2} N_2^{2d-2}$$

This example shows that when $d \geq 3$, the term with N_2^{d-3}/λ is sharp.

Acknowledgment

We thank Larry Guth for very helpful discussions during the course of this work.

References

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

[Bourgain 1998] J. Bourgain, "Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity", *Internat. Math. Res. Notices* **1998**:5 (1998), 253–283. MR Zbl

[Bourgain 2013] J. Bourgain, "Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces", *Israel J. Math.* **193**:1 (2013), 441–458. MR Zbl

[Bourgain and Demeter 2015] J. Bourgain and C. Demeter, "The proof of the l^2 decoupling conjecture", *Ann. of Math.* (2) **182**:1 (2015), 351–389. MR Zbl

[Bourgain and Demeter 2017] J. Bourgain and C. Demeter, "A study guide for the l^2 decoupling theorem", *Chin. Ann. Math. Ser. B* **38**:1 (2017), 173–200. MR Zbl

[Colliander et al. 2002] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, "Almost conservation laws and global rough solutions to a nonlinear Schrödinger equation", *Math. Res. Lett.* **9**:5 (2002), 659–682. MR Zbl

[De Silva et al. 2007] D. De Silva, N. Pavlović, G. Staffilani, and N. Tzirakis, "Global well-posedness for a periodic nonlinear Schrödinger equation in 1D and 2D", *Discrete Contin. Dyn. Syst.* **19**:1 (2007), 37–65. MR Zbl

[Deng and Germain 2017] Y. Deng and P. Germain, "Growth of solution of NLS on irrational tori", preprint, 2017. arXiv

[Deng et al. 2017] Y. Deng, P. Germain, and L. Guth, "Strichartz estimates for the Schrödinger equation on irrational tori", *J. Funct. Anal.* **273**:9 (2017), 2846–2869. MR Zbl

[Guo et al. 2014] Z. Guo, T. Oh, and Y. Wang, "Strichartz estimates for Schrödinger equations on irrational tori", *Proc. Lond. Math. Soc.* (3) **109**:4 (2014), 975–1013. MR Zbl

[Hani 2012] Z. Hani, "A bilinear oscillatory integral estimate and bilinear refinements to Strichartz estimates on closed manifolds", *Anal. PDE* **5**:2 (2012), 339–363. MR Zbl

[Herr et al. 2011] S. Herr, D. Tataru, and N. Tzvetkov, "Global well-posedness of the energy-critical nonlinear Schrödinger equation with small initial data in $H^1(\mathbb{T}^3)$", *Duke Math. J.* **159**:2 (2011), 329–349. MR Zbl

[Ionescu and Pausader 2012] A. D. Ionescu and B. Pausader, "The energy-critical defocusing NLS on $\mathbb{T}^3$", *Duke Math. J.* **161**:8 (2012), 1581–1612. MR Zbl

[Killip and Vişan 2016] R. Killip and M. Vişan, "Scale invariant Strichartz estimates on tori and applications", *Math. Res. Lett.* **23**:2 (2016), 445–472. MR Zbl

[Ramos 2016] J. Ramos, "The trilinear restriction estimate with sharp dependence on the transversality", preprint, 2016. arXiv

Received 23 Jan 2017. Revised 7 Oct 2017. Accepted 11 Dec 2017.

CHENJIE FAN: `cjfan@math.uchicago.edu`
Department of Mathematics, University of Chicago, Chicago, IL, United States

GIGLIOLA STAFFILANI: `gigliola@math.mit.edu`
Department of Mathematics, Massachusetts Institue of Technology, Cambridge, MA, United States

HONG WANG: `hongwang@mit.edu`
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

BOBBY WILSON: `blwilson@mit.edu`
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, United States

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2018 is US $275/year for the electronic version, and $480/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

Analysis & PDE

Volume 11 No. 4 2018

C^1 regularity of orthotropic p-harmonic functions in the plane 813
PIERRE BOUSQUET and LORENZO BRASCO

Applications of small-scale quantum ergodicity in nodal sets 855
HAMID HEZARI

On rank-2 Toda systems with arbitrary singularities: local mass and new estimates 873
CHANG-SHOU LIN, JUN-CHENG WEI, WEN YANG and LEI ZHANG

Beyond the BKM criterion for the 2D resistive magnetohydrodynamic equations 899
LÉO AGÉLAS

On a bilinear Strichartz estimate on irrational tori 919
CHENJIE FAN, GIGLIOLA STAFFILANI, HONG WANG and BOBBY WILSON

Sharp global estimates for local and nonlocal porous medium-type equations in bounded domains 945
MATTEO BONFORTE, ALESSIO FIGALLI and JUAN LUIS VÁZQUEZ

Blow-up of a critical Sobolev norm for energy-subcritical and energy-supercritical wave equations 983
THOMAS DUYCKAERTS and JIANWEI YANG

Global weak solutions for generalized SQG in bounded domains 1029
HUY QUANG NGUYEN

Scale-free unique continuation principle for spectral projectors, eigenvalue-lifting and Wegner estimates for random Schrödinger operators 1049
IVICA NAKIĆ, MATTHIAS TÄUFER, MARTIN TAUTENHAHN and IVAN VESELIĆ

2157-5045(2018)11:4;1-Q