R. Fraser *et al.* (2022) "Polynomial Roth Theorems on Sets of Fractional Dimensions,"
International Mathematics Research Notices, Vol. 2022, No. 10, pp. 7809–7838
Advance Access Publication January 6, 2021
https://doi.org/10.1093/imrn/rnaa377



# Polynomial Roth Theorems on Sets of Fractional Dimensions

**Robert Fraser[1], Shaoming Guo[2,*] and Malabika Pramanik[3]**

[1]School of Mathematics, University of Edinburgh, Edinburgh EH9 3FD, UK, [2]Department of Mathematics, University of Wisconsin Madison, 53706, USA, and [3]Department of Mathematics, University of British Columbia, Vancouver, V6T 1Z2, Canada

**Correspondence to be sent to: e-mail: shaomingguo2018@gmail.com*

Let $E \subset \mathbb{R}$ be a closed set of Hausdorff dimension $\alpha \in (0,1)$. Let $P : \mathbb{R} \to \mathbb{R}$ be a polynomial without a constant term whose degree is bigger than one. We prove that if $E$ supports a probability measure satisfying certain dimension condition and Fourier decay condition, then $E$ contains three points $x, x+t, x+P(t)$ for some $t > 0$. Our result extends the one of Łaba and Pramanik [11] to the polynomial setting, under the same assumption. It also gives an affirmative answer to a question in Henriot *et al.* [7].

## 1 Statement of Results

This paper is dedicated to the following question: when does a set $E \subseteq [0,1]$ contain a "polynomial configuration"? More precisely, given a univariate real polynomial $P$, under what condition on $E$ can we ensure the existence of a triple of distinct points $(x_1, x_2, x_3) \in E^3$ satisfying $x_3 - x_1 = P(x_2 - x_1)$? We address this question through measure-theoretic assumptions on $E$. Specifically, we assume that $E$ supports a probability measure $\mu$ that obeys two conditions; the 1st is a ball condition of order $\alpha$ and the 2nd is a Fourier decay condition of order $\beta$. These are specified below:

$$\begin{aligned} (A)_\alpha \qquad & \sup_{\epsilon \in (0,1]} \epsilon^{-\alpha} \mu([x, x+\epsilon]) \le C_1, \\ (B)_\beta \qquad & \sup_{k \in \mathbb{Z}} (1+|k|)^{\frac{\beta}{2}} |\widehat{\mu}(k)| \le C_2 (1-\alpha)^{-B}. \end{aligned} \tag{1.1}$$

Received November 19, 2019; Revised October 26, 2020; Accepted December 1, 2020
Communicated by Prof. Tim Gowers

Here, $\widehat{\mu}$ denotes the Fourier transform of the measure $\mu$:

$$\widehat{\mu}(\xi) := \int e^{-ix\xi}\, \mathrm{d}\mu(x), \qquad \xi \in \mathbb{R}.$$

Our main result is that the confluence of the two conditions in (1.1) for appropriate choices of $\alpha, \beta$ is sufficient to ensure polynomial configurations.

**Theorem 1.1.** For every polynomial $P : \mathbb{R} \to \mathbb{R}$ of degree at least 2 with $P(0) = 0$, one can find a constant $s_0 > 0$ depending only on $P$ such that the following statement holds. For every choice of positive constants $C_1, C_2, B$, there exists $\alpha_0 \in (0, 1)$ depending only on $s_0$ and these constants such that any compact set $E \subset [0, 1]$ that supports a probability measure $\mu$ satisfying the relations $(A)_\alpha$ and $(B)_\beta$ in (1.1) also contains three points

$$x, x + t, x + P(t) \text{ for some } t > 0, \tag{1.2}$$

provided $\alpha \in (\alpha_0, 1)$ and $\beta \in (1 - s_0, 1)$.

Our proof shows that the value of $t$ in the above theorem is extremely small. When $P(t)$ does not contain any constant term, the value of $P(t)$ will also be small. In this sense, our problem is local. The problem of dealing with a polynomial $P(t)$ containing a constant term is more global and is not covered here. A version of Theorem 1.1 with $P(t) = 2t$ is due to Łaba and Pramanik [11], with $s_0 = 1/3$. The assumptions $(A)_\alpha$ and $(B)_\beta$ first appeared here. These assumptions turn out to be very natural in the context of the sets of large Fourier dimension and also Salem sets. Indeed, many probabilistic constructions that yield measures obeying $(A)_\alpha$ and $(B)_\beta$ also produce Salem sets. However, it is important to note that not every Salem set obeys $(A)_\alpha$ and $(B)_\beta$; see for instance [15].

Let us briefly recall the discussion in [11]. For a Borel set $E \subset [0, 1]$, we let $\dim_{\mathcal{H}}(E)$ denote the Hausdorff dimension of the set $E$. It is well known [14] that $\dim_{\mathcal{H}}(E)$ is the supremum of numbers $\alpha \in [0, 1]$ for which there exists a probability measure $\mu$ supported on $E$ satisfying

$$\sup \left\{ r^{-\alpha} \mu\left(B(x; r)\right) : x \in \mathbb{R},\ r > 0 \right\} < \infty.$$

In contrast, the *Fourier dimension* of $E$, denoted $\dim_{\mathcal{F}}(E)$, is defined to be the supremum overall $\beta \in [0, 1]$ such that there exists a probability measure $\mu$ supported on $E$ satisfying

$$|\widehat{\mu}(\xi)| \le C(1 + |\xi|)^{-\beta/2} \text{ for every } \xi \in \mathbb{R} \text{ and some } C > 0. \tag{1.3}$$

Thus, the condition $(A)_\alpha$ implies that $E$ has Hausdorff dimension at least $\alpha$, while the condition $(B)_\beta$ implies that $E$ has Fourier dimension at least $\beta$. It is known [14] that

$$\dim_{\mathcal{F}}(E) \leq \dim_{\mathcal{H}}(E) \text{ for every } E. \tag{1.4}$$

The sets for which the equality in (1.4) is achieved are called Salem sets. So far, there are a number of constructions of Salem sets, due to Salem [17], Kaufman [10] (see Bluhm [2] for an exposition), Kahane [9], Bluhm [1], Łaba and Pramanik [11], and so on. Many of these constructions are probabilistic constructions. For instance, Kahane [9] showed that images of compact sets under Brownian motion are almost surely Salem sets.

It is worth mentioning that, in Salem's probabilistic construction of Salem [17] sets, with large probability, the examples there (under certain modifications as in [11]) obey assumptions $(A)$ and $(B)$. Moreover, Łaba and Pramanik [11] also provided a probabilistic construction of Salem sets, a large portion of which (under a natural measure) satisfy assumptions $(A)$ and $(B)$.

We will discuss a few generalizations of the result of Łaba and Pramanik [11]. In [4], Chan, Łaba, and Pramanik generalized [11] to higher dimensions. Their result covers a large class of linear patterns. In particular, they proved the following: let $a, b, c$ be three points in the plane that are not co-linear. Let $E \subset \mathbb{R}^2$. Assume that $E$ supports a probability measure $\mu$ satisfying the analogues of assumptions $(A)$ and $(B)$ in $\mathbb{R}^2$. Then, $E$ must contain three distinct points $x, y, z$ such that the triangle $\Delta xyz$ is similar to $\Delta abc$.

The result of [4] was later generalized to certain nonlinear patterns by Henriot *et al.* [7]. However, their result does not cover the case of dimension one. For instance, it was pointed out by the authors of [7] that the configuration $(x, x+t, x+t^2)$ with $x, t \in \mathbb{R}$ cannot be detected by their method. In the current paper, we provide an affirmative answer to this question.

Before we describe the proof of the main theorem, let us also mention that results on the reals that are analogous to Theorem 1.1 have been obtained by Bourgain [3] and Durcik *et al.* [5]. Indeed, some of the techniques we use also originate from these works. This will be discussed after stating Proposition 1.2, one main ingredient of the proof of Theorem 1.1.

We turn to the proof of Theorem 1.1. Let $\tau_0$ be a nonnegative smooth bump function supported on the interval $[1, 2]$ and $\tau_l(t) := \tau_0(2^l t)$. It is not difficult to imagine that the trilinear form

$$\iint \mu(x)\mu(x+t)\mu(x+P(t))\tau_l(t)\, \mathrm{d}t\, \mathrm{d}x \tag{1.5}$$

will play a crucial role in the study of patterns as in our main theorem. However, as $\mu$ is just a measure, the above trilinear form may not be well defined at the 1st place. Our 1st task is to make sense of this trilinear form for every integer $l$ that is large enough.

Let $s$ be a real number. Define a Sobolev norm

$$\|f\|_{H^s} := \left( \int_{\mathbb{R}} |\hat{f}(\xi)|^2 (1 + |\xi|^2)^s \, \mathrm{d}\xi \right)^{1/2}. \tag{1.6}$$

For $l \in \mathbb{N}$, and two Schwartz functions $f$ and $g$, define

$$T_l(f, g)(x) := \int_{\mathbb{R}} f(x+t) g(x + P(t)) \tau_l(t) \, \mathrm{d}t. \tag{1.7}$$

We will prove the following proposition.

**Proposition 1.2.** There exists a small constant $s_0 > 0$ and large constant $l_0 > 0$ and $\gamma_0 > 0$, depending only on $P(t)$, such that

$$\|T_l(f, g)\|_{H^{s_0}} \leq 2^{\gamma_0 l} \|f\|_{H^{-s_0}} \|g\|_{H^{-s_0}}, \tag{1.8}$$

for every $l \geq l_0$, and for Schwartz functions $f$ and $g$.

This is called a Sobolev improving estimate. To our knowledge, an estimate of this form first appeared in the work of Bourgain [3]; see Lemma 5 there. Li [12] and Lie [13] further developed it in the context of Hilbert transforms along the curves. All these works require that the polynomial $P(t)$ does not have a linear term, in order to use a certain "curvature" property of the polynomial. In Proposition 1.2, we managed to prove a Sobolev improving estimate for every polynomial. The proof is a variant of Li [12]. Moreover, it does not require the notion of $\sigma$-uniformity there. One reason that [12] requires $P(t)$ to have no linear terms is that certain curvature condition (nonvanishing of the left-hand side of (2.48)) would fail for polynomials with linear terms. The key (and simple) observation in the current paper is that the curvature condition fails only on a "small" set (see Claim 2.5).

After proving Proposition 1.2, we are able to use it to make sense of the double integral in (1.5). Let $\mu$ be a probability measure supported on the interval $[0, 1]$. If we also assume that

$$|\hat{\mu}(\xi)| \leq C|\xi|^{-\frac{\beta}{2}}, \text{for} \beta \in (1 - s_0, 1), \tag{1.9}$$

and for some constant $C > 0$, then $\mu \in H^{-s_0}(\mathbb{R})$, which is a Sobolev space of some negative order. Recall that Schwartz functions are dense in $H^s$ for every $s \in \mathbb{R}$. By a density argument, we know that the double integral in (1.5) is well defined. To be precise, we will pick a sequence of Schwartz functions $\{f_n\}_{n=1}^{\infty}$ that convergence to $\mu$ in $H^{-s_0}$ and interpret (1.5) as

$$\lim_{n\to\infty} \iint f_n(x) f_n(x+t) f_n(x+P(t)) \tau_l(t)\, \mathrm{d}t\, \mathrm{d}x. \tag{1.10}$$

That the above limit exists is guaranteed by Proposition 1.2.

After making sense of the double integral in (1.5), we will prove that it is always positive. That is, we will prove the following theorem.

**Theorem 1.3.** Under the same assumptions as in Theorem 1.1, we are able to find a large integer $l_0 \in \mathbb{N}$ and a small positive real number $c_0 > 0$ such that

$$\iint \mu(x)\mu(x+t)\mu(x+P(t))\tau_{l_0}(t)\, \mathrm{d}t\, \mathrm{d}x \geq c_0. \tag{1.11}$$

Intuitively speaking, if $E$ does not contain any three-term configuration $(x, x+t, x+P(t))$, then the left-hand side of (1.11) would certainly vanish. However, as we deal with the measures supported on the sets of fractional dimensions, we need some extra work to make the above argument rigorous. Roughly speaking, we will construct a Borel measure $\nu$ defined on $[0,1]^2$ and supported on the set of configurations $(x, x+t, x+P(t))$ with $t > 0$, such that $\nu([0,1]^2) > 0$. This will guarantee the existence of the desired polynomial pattern. This will be carried out in the last section.

**Organization of the paper.** The Sobolev improving estimate in Proposition 1.2 will be proven in Section 2. The main tools we will be using include the stationary phase principle and techniques from bilinear oscillatory integrals recently developed by Li [12]. In Section 3, we provide a proof of the stationary phase principle that is used in the current paper. Theorem 1.3 will be proven in Sections 4 and 5. The argument that is used in this step relies on the idea of measure decomposition of Łaba and Pramanik [11], on the Sobolev improving estimate in Proposition 1.2 and on Bourgain's energy pigeonholing argument from [3]. Finally, in Section 6, we will finish the proof of Theorem 1.1.

**Notation.** Throughout the paper, we will write $x \lesssim y$ to mean that there exists a universal constant $C$ such that $x \leq Cy$ and $x \simeq y$ to mean that $x \lesssim y$ and $y \lesssim x$. Moreover,

$x \lesssim_{M,N} y$ means there exists a constant $C_{M,N}$ depending on the parameters $M$ and $N$ such that $x \le C_{M,N} y$.

## 2 Sobolev Improving Estimate: Proof of Proposition 1.2

In this section, we prove Proposition 1.2. Let $P : \mathbb{R} \to \mathbb{R}$ be a polynomial of degree bigger than one without constant term. We write it as

$$P(t) = a_n t^{\alpha_n} + \cdots + a_2 t^{\alpha_2} + a_1 t^{\alpha_1}, \tag{2.1}$$

with $1 \le \alpha_1 < \alpha_2 < \cdots < \alpha_n$. Here, we assume that $a_i \neq 0$ for every $i \in \{1, 2, \ldots, n\}$. Moreover, we assume that $\alpha_1 = 1$, that is, our polynomial $P$ contains a linear term. The corresponding result for a polynomial without linear term is much easier to prove. This point will be elaborated in a few lines.

For each $1 \le i \le n$, let $b_i$ be the unique integer such that

$$2^{b_i} \le |a_i| < 2^{b_i+1}. \tag{2.2}$$

Let $\Gamma_0$ be a large number that depends on the polynomial $P$. Let $l_0 \in \mathbb{N}$ be the smallest integer such that for every $l \ge l_0$, the following hold:

$$|a_i| 2^{-l\cdot\alpha_i} \le \Gamma_0^{-1} |a_1 2^{-\alpha_1 l}| \text{ for every } n \ge i \ge 2 \tag{2.3}$$

and

$$|a_i| 2^{-l\cdot\alpha_i} \le \Gamma_0^{-1} |a_2 2^{-l\cdot\alpha_2}| \text{ for every } n \ge i \ge 3. \tag{2.4}$$

In other words, at the scale $t \simeq 2^{-l}$, the monomial $a_1 t$ "dominates" the polynomial $P(t)$ and $a_2 t^{\alpha_2}$ is the 2nd dominating term. It is not difficult to see that the choice of $l_0$ depends only on $P(t)$.

Let us pause and make a remark on the assumption that $\alpha_1 = 1$. As mentioned above, the case $\alpha_1 > 1$ is relatively easier to handle. This is because a certain curvature (in the sense of oscillatory integrals) appears naturally in this case. To be more precise, under the assumption that $\alpha_1 \ge 2$, we first choose $l$ large enough such that (2.3) holds and then the monomial $a_1 t^{\alpha_1}$ dominates the polynomial $P(t)$ at the scale $t \simeq 2^{-l}$. Notice that acertain curvature is already present when $a_1 t^{\alpha_1}$ dominates. Hence, the requirement (2.4) becomes redundant.

However, under the assumption that $\alpha_1 = 1$, if we only require (2.3), then there is no curvature in the dominating term $a_1 t$. This is why we need to further require (2.4) and find a 2nd dominating term. It is hoped that the curvature in the 2nd dominating term will play an equivalent role. Due to the presence of the linear term $a_1 t$, a number of extra complications will appear.

The rest of this section is devoted to the proof of Proposition 1.2. Let $h$ be a function in $H^{-s_0}$. We pair it with the left-hand side of (1.8) and study

$$\int_{\mathbb{R}} \left[ \int_{\mathbb{R}} f(x+t) g(x+P(t)) \tau_l(t) \, \mathrm{d}t \right] h(x) \, \mathrm{d}x. \tag{2.5}$$

Let $\psi_0 : \mathbb{R} \to \mathbb{R}$ be a nonnegative smooth bump function supported on $[-3,-1] \cup [1,3]$. Define $\psi_k(\cdot) = \psi_0(\cdot/2^k)$. Moreover, we choose $\psi_0$ such that

$$1 = \sum_{k \in \mathbb{Z}} \psi_k(t), \text{ for every } t \neq 0. \tag{2.6}$$

For all the three functions $f, g$, and $h$, we apply the nonhomogeneous Littlewood–Paley decomposition $\mathbb{1} = \sum_{k \in \mathbb{N}} P_k$, where $\mathbb{1}$ denotes the identity operator, and study

$$\sum_{k_1,k_2,k_3=0}^{\infty} \iint_{\mathbb{R}^2} P_{k_1} f(x+t) P_{k_2} g(x+P(t)) \tau_l(t) P_{k_3} h(x) \, \mathrm{d}t \, \mathrm{d}x. \tag{2.7}$$

Here,

$$P_k f(x) := \int_{\mathbb{R}} e^{ix\xi} \psi_k(\xi) \hat{f}(\xi) \, \mathrm{d}\xi, \text{if } k > 0 \tag{2.8}$$

and

$$P_0 f(x) := \int_{\mathbb{R}} e^{ix\xi} \left( \sum_{k \leq 0} \psi_k(\xi) \right) \hat{f}(\xi) \, \mathrm{d}\xi. \tag{2.9}$$

In the following, we work on two cases,

$$|(k_1 - l) - (k_2 - l + b_1)| \geq 100 \text{ and } |(k_1 - l) - (k_2 - l + b_1)| < 100. \tag{2.10}$$

Let us begin with the 1st case. Our goal is to prove the following lemma.

**Lemma 2.1.** There exists a constant $\gamma_0 \in \mathbb{N}$ depending on $P$, such that under the assumption that $|k_1 - k_2 - b_1| \geq 100$, we have

$$\left| \iint_{\mathbb{R}^2} P_{k_1} f(x+t) P_{k_2} g(x+P(t)) \tau_l(t) P_{k_3} h(x) \, \mathrm{d}t \, \mathrm{d}x \right| \lesssim_N 2^{\gamma_0 l} 2^{-N(k_1+k_2+k_3)} \|P_{k_1} f\|_2 \|P_{k_2} g\|_2 \|P_{k_3} h\|_2, \tag{2.11}$$

for arbitrarily large $N \in \mathbb{N}$.

Assuming the above lemma, we have

$$\sum_{\substack{k_1,k_2,k_3 \\ |k_1-k_2-b_1|\geq 100}} \left| \iint_{\mathbb{R}^2} P_{k_1} f(x+t) P_{k_2} g(x+P(t)) \tau_l(t) P_{k_3} h(x) \, \mathrm{d}t \, \mathrm{d}x \right| \lesssim_l \sum_{k_1,k_2,k_3=0}^{\infty} 2^{-10(k_1+k_2+k_3)} \|P_{k_1} f\|_2 \|P_{k_2} g\|_2 \|P_{k_3} h\|_2 \lesssim_l \|f\|_{H^{-s_0}} \|g\|_{H^{-s_0}} \|h\|_{H^{-s_0}}, \tag{2.12}$$

for some $s_0 > 0$. For instance, we may take $s_0 = 1$ at this step.

**Proof of Lemma 2.1.** The proof is via an integration by parts. Turning to the Fourier side, we can write the left-hand side of (2.11) as

$$2^{-l} \left| \iint \widehat{P_{k_1} f}(\xi) \widehat{P_{k_2} g}(\eta) \widehat{P_{k_3} h}(\xi+\eta) \left[ \int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta} \tau_0(t) \, \mathrm{d}t \right] \mathrm{d}\xi \, \mathrm{d}\eta \right|. \tag{2.13}$$

First of all, we observe that

$$(2.13) = 0 \text{ when } k_3 \geq k_1 + k_2 + 10. \tag{2.14}$$

Hence, in the rest of the proof, we assume that $k_3 \leq k_1 + k_2 + 10$, and it suffices to prove

$$(2.13) \lesssim_l 2^{-N(k_1+k_2)} \|P_{k_1} f\|_2 \|P_{k_2} g\|_2 \|P_{k_3} h\|_2. \tag{2.15}$$

Here, $N \in \mathbb{N}$ is a large integer that might vary from line to line. By an integration by parts, we obtain

$$\left| \int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta} \tau_0(t) \, \mathrm{d}t \right| \lesssim_l 2^{-N \max\{k_1,k_2\}}, \tag{2.16}$$

for $\xi$ and $\eta$ in the frequency support of $P_{k_1}f$ and $P_{k_2}g$, respectively. Substitute the above pointwise bound into (2.13), and apply Hölder's inequality in the $\xi$ and $\eta$ variables. This will finish the proof of the desired estimate. It is easy to track the dependence on $l$ and see that it is polynomial in $2^l$. ■

From now on, we may assume that $|k_1-k_2-b_1| < 100$. Without loss of generality, we take $k_1 = k_2 + b_1$ and consider

$$\sum_{k_1,k_3=0}^{\infty} \iint_{\mathbb{R}^2} P_{k_1}f(x+t)P_{k_1-b_1}g(x+P(t))\tau_l(t)P_{k_3}h(x)\,\mathrm{d}t\,\mathrm{d}x. \tag{2.17}$$

In the double sum over $k_1$ and $k_3$, we may impose the extra condition that $k_3 \le 2(k_1 + |b_1|)$, as otherwise the corresponding term from (2.17) will simply vanish.

**Lemma 2.2.** There exists a constant $\gamma_0 \in \mathbb{N}$ and $\gamma > 0$, both of which are allowed to depend on $P$, such that

$$\left|\iint_{\mathbb{R}^2} f(x+t)g(x+P(t))\tau_l(t)h(x)\,\mathrm{d}t\,\mathrm{d}x\right| \le 2^{\gamma_0 l}2^{-\gamma k}\|f\|_2\|g\|_2\|h\|_2, \tag{2.18}$$

for every $l \ge l_0$ and every $k \in \mathbb{N}$, under the assumption that $\mathrm{supp}(\hat{f}) \subset \pm[2^k, 2^{k+1}]$ and $\mathrm{supp}(\hat{g}) \subset \pm[2^{k-b_1}, 2^{k-b_1+1}]$ and no further assumption on the function $h$.

Assuming this lemma, we will be able to finish the proof of the desired bilinear estimate. Recall that we need to control (2.17). By Lemma 2.2, this can be bounded by

$$2^{\gamma_0 l} \sum_{k_1,k_3 \text{ with } k_3 \le 2k_1+2|b_1|} 2^{-\gamma k_1}\|P_{k_1}f\|_2\|P_{k_1-b_1}g\|_2\|P_{k_3}h\|_2, \tag{2.19}$$

for some $\gamma > 0$, which can be further bounded by

$$\|f\|_{H^{-\gamma/6}}\|g\|_{H^{-\gamma/6}}\|h\|_{H^{-\gamma/6}}. \tag{2.20}$$

This finishes the proof of the desired estimate.

Hence, it remains to prove Lemma 2.2. As the constant is allowed to depend on $l$, we can always assume that $k$ is at least some large constant times $\ell$. Turning to the Fourier side, we obtain

$$2^{-l}\iint \widehat{P_k f}(\xi)\widehat{P_{k-b_1}g}(\eta)\hat{h}(\xi+\eta)\left[\int_{\mathbb{R}} e^{i2^{-l}t\xi+iP(2^{-l}t)\eta}\tau_0(t)\,\mathrm{d}t\right]\mathrm{d}\xi\,\mathrm{d}\eta. \tag{2.21}$$

Write

$$P(t) = a_1 t + Q(t) = a_1 t + a_2 t^{\alpha_2} + R(t). \tag{2.22}$$

The derivative of the phase function in (2.21) is given by

$$2^{-l}\xi + a_1 2^{-l}\eta + a_2\alpha_2 2^{-\alpha_2 l} t^{\alpha_2 - 1}\eta + 2^{-l}R'(2^{-l}t)\eta. \tag{2.23}$$

From the 1st-order derivative of the phase function, we are still not able to locate the critical point. To do so, we apply a more refined frequency decomposition to $f$ and $g$. For a fixed integer $\Delta$, let $\psi_{k,l,\Delta} : \mathbb{R} \to \mathbb{R}$ be a nonnegative smooth function supported on $[2^k + \Delta \cdot 2^{k-\gamma_0 l}, 2^k + (\Delta+2)2^{k-\gamma_0 l}]$ such that

$$\psi_k(\xi) = \sum_{\Delta\in\mathbb{Z}} \psi_k(\xi)\psi_{k,l,\Delta}(\xi), \text{ for every } \xi \in \mathbb{R}. \tag{2.24}$$

That is, $\{\psi_{k,l,\Delta}\}_{\Delta\in\mathbb{Z}}$ forms a partition of unity on the support of $\psi_k$. Moreover, the sum in (2.24) is indeed a finite sum and the number of nonzero terms is about $2^{\gamma_0 l}$. Here, $\gamma_0$ is some large number that is to be chosen. For convenience, we will allow $\gamma_0$ to change from line to line, unless otherwise stated.

We write (2.21) as

$$\sum_{\Delta_1,\Delta_2\in\mathbb{Z}} \iint \left[\widehat{P_k f}(\xi)\psi_{k,l,\Delta_1}(\xi)\right]\left[\widehat{P_{k-b_1} g}(\eta)\psi_{k-b_1,l,\Delta_2}(\eta)\right]\hat{h}(\xi+\eta)\left[\int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta}\tau_0(t)\,\mathrm{d}t\right]\mathrm{d}\xi\,\mathrm{d}\eta. \tag{2.25}$$

Notice that in the above sum, we have about $2^{2\gamma_0 l}$ terms that may be nonzero. We will see that the main contribution to the sum will come from those $\Delta_1, \Delta_2$ such that the phase

$$2^{-l}t\xi + P(2^{-l}t)\eta$$

admits a critical point $t \in [1, 2]$ and some $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$, $\eta \in \operatorname{supp}\psi_{k,l,\Delta_2}$.

Observe that the critical points of the phase are those values of $t$ such that

$$2^{-l}\xi + a_1 2^{-l}\eta + a_2\alpha_2 2^{-\alpha_2 l} t^{\alpha_2-1}\eta + 2^{-l}R'(2^{-l}t)\eta = 0.$$

We have, by (2.4), that for $t \in [1, 2]$, we have $|R'(2^{-l}t)| \le 2^{\alpha_n}\alpha_n\Gamma_0^{-1}|a_2|2^{-(\alpha_2-1)l}$. Thus, if $t$ is a critical point of the phase function, we must have

$$|2^{-l}\xi + a_1 2^{-l}\eta + a_2\alpha_2 2^{-\alpha_2 l} t^{\alpha_2-1}\eta| \le 2^{\alpha_n}\alpha_n\Gamma_0^{-1}|a_2|2^{-\alpha_2 l}\eta. \tag{2.26}$$

Thus, we get by dividing that

$$\left|\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} + t^{\alpha_2 - 1}\right| \le \frac{2^{\alpha_n}\alpha_n \Gamma_0^{-1}|a_2|}{a_2\alpha_2}. \tag{2.27}$$

Observe that, as long as $l$ is sufficiently large that (2.4) is satisfied, the right side of this inequality does not depend on $l$. So, if $\Gamma_0$ is sufficiently large depending on the polynomial $P$, then the right-hand side of (2.27) is bounded above by $\frac{1}{100}$. Thus, if $\xi$ and $\eta$ are such the phase function has a critical point in $[1, 2]$, we must have the inequality

$$\frac{99}{100} \le -\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} \le 2^{\alpha_2 - 1} + \frac{1}{100}. \tag{2.28}$$

Now, if (2.28) holds for some $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$ and $\eta \in \operatorname{supp}\psi_{k-b_1,l,\Delta_2}$, then by the mean-value theorem, we have that if $l$ is sufficiently large depending on the polynomial $P$ and the parameter $\gamma_0$, then

$$\frac{1}{2} \le -\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} \le 2^{\alpha_2 - 1} + \frac{1}{2} \tag{2.29}$$

for *all* such $\xi$ and $\eta$.

We then partition into two sums $S_{\text{crit}} + S_{\text{err}}$, where $S_{\text{crit}}$ is the sum over those $\Delta_1, \Delta_2$ such that (2.29) holds for all $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$ and $\eta \in \operatorname{supp}\psi_{k-b_1,l,\Delta_2}$, and $S_{\text{err}}$ is the sum over those $\Delta_1, \Delta_2$ such that (2.29) fails for some $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$ and $\eta \in \operatorname{supp}\psi_{k-b_1,l,\Delta_2}$.

We will first estimate $S_{\text{err}}$.

**Lemma 2.3** (Nonstationary phase estimate). For $S_{\text{err}}$ as defined above, we have

$$|S_{\text{err}}| \lesssim_l 2^{-k/2}||\widehat{P_k f}||_2||\widehat{P_{k-b_1} g}||_2||\hat{h}||_2 \tag{2.30}$$

and the dependence is polynomial in $2^l$.

**Proof.** We will only estimate the contribution from the terms for which $-\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} > 2^{\alpha_2 - 1} + \frac{1}{2}$ for some $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$ and $\eta \in \operatorname{supp}\psi_{k-b_1,l,\Delta_2}$; the terms for which $-\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} < \frac{1}{2}$ are handled similarly.

If $l$ is sufficiently large, an argument similar to the one establishing (2.29) from (2.28) shows that

$$-\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} > 2^{\alpha_2 - 1} + \frac{1}{4} \tag{2.31}$$

for *all* $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$ and $\eta \in \operatorname{supp}\psi_{k-b_1,l,\Delta_2}$.

We will now reverse the argument used to infer (2.28) from (2.26) in order to show that the derivative of the phase is large for $t \in [1,2]$. To start, we subtract $t^{\alpha_2-1}$ from both sides of (2.31) and take absolute values. Since $t^{\alpha_2-1} \leq 2^{\alpha_2-1}$, we get

$$\left|\frac{2^{-l}\xi + a_1 2^{-l}\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} + t^{\alpha_2-1}\right| > \frac{1}{4}. \tag{2.32}$$

Because $l$ is such that $\frac{2^{-l}R'(2^{-l}t)\eta}{a_2\alpha_2 2^{-\alpha_2 l}\eta} < \frac{1}{100}$, we have

$$\left|2^{-l}\xi + a_1 2^{-l}\eta + a_2\alpha_2 2^{-\alpha_2 l}t^{\alpha_2-1}\eta + 2^{-l}R'(2^{-l}t)\eta\right| > \frac{24}{100}a_2\alpha_2 2^{-\alpha_2 l}\eta. \tag{2.33}$$

The left side of (2.33) is the derivative of the phase function and the right side of (2.33) is $\gtrsim 2^{k-\alpha_2 l}$, where the implicit constant depends only on the polynomial $P$. Thus, integrating by parts once gives that for $\xi \in \operatorname{supp}\psi_{k,l,\Delta_1}$, $\eta \in \operatorname{supp}\psi_{k,l,\Delta_2}$, we have the estimate

$$\left|\int e^{i2^{-l}t\xi + iP(2^{-l}t)\eta}\tau_0(t)\,\mathrm{d}t\right| \lesssim 2^{-k+\alpha_2 l}. \tag{2.34}$$

Because there are only at most $2^{2\gamma_0 l}$ terms that contribute to the sum, we have from applying the Cauchy–Schwarz inequality and observing that $\widehat{P_k f}$ is supported on a set of measure $\lesssim 2^k$,

$$|S_{\mathrm{err}}| \lesssim 2^{-k/2+(2\gamma_0+\alpha_2)l}||\widehat{P_k f}||_2||\widehat{P_{k-b_1}g}||_2||\hat{h}||_2, \tag{2.35}$$

as desired. ■

We now estimate $S_{\mathrm{crit}}$. As the implicit constant is allowed to depend on $l$, and there are only at most $2^{\gamma_0 l}$ values of $\Delta_1$ and $\Delta_2$ that contribute to (2.4), it suffices to bound each term of $S_{\mathrm{crit}}$ separately. After this reduction, what we need to prove becomes

$$\left|\iint \hat{f}(\xi)\hat{g}(\eta)\hat{h}(\xi+\eta)\left[\int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta}\tau_0(t)\,\mathrm{d}t\right]\mathrm{d}\xi\,\mathrm{d}\eta\right| \leq 2^{\gamma_0 l}2^{-\gamma k}\|f\|_2\|g\|_2\|h\|_2, \tag{2.36}$$

under the assumption that

$$\hat{f} = \widehat{P_k f} \cdot \psi_{k,l,\Delta_1}, \quad \hat{g} = \widehat{P_{k-b_1} g} \cdot \psi_{k-b_1,l,\Delta_2} \tag{2.37}$$

and that (2.29) holds for every $\xi \in \operatorname{supp}(\hat{f})$ and $\eta \in \operatorname{supp}(\hat{g})$. Observe that if (2.29) holds for a given $\xi$ and $\eta$, the critical point $t_c$ of the phase function must lie in $\left(\frac{1}{4}, 4\right)$ provided that $\Gamma_0$ is sufficiently large by (2.27).

The critical point $t_c$ is given by

$$\xi + P'(2^{-l}t_c)\eta = 0. \tag{2.38}$$

We will prove the following approximation formula.

**Lemma 2.4** (Method of stationary phase). Under the above notation, we have

$$\int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta} \tau_0(t)\, \mathrm{d}t = a(\xi,\eta)\eta^{-1/2} e^{i\Psi(\xi,\eta)} + O_l\left(\frac{1}{|\eta|}\right), \tag{2.39}$$

with

$$a(\xi,\eta) := (2^{-2l}P''(2^{-l}t_c))^{-1/2}\tau_0(t_c) \tag{2.40}$$

and

$$\Psi(\xi,\eta) := 2^{-l}t_c\xi + P(2^{-l}t_c)\eta. \tag{2.41}$$

$\Psi(\xi,\eta)$ satisfies the equation

$$\partial_\xi \Psi(\xi,\eta) = 2^{-l}t_c. \tag{2.42}$$

Moreover,

$$O_l\left(\frac{1}{|\eta|}\right) \leq 2^{\gamma_0 l}\frac{1}{|\eta|}. \tag{2.43}$$

Observe that $a(\xi,\eta) = 0$ unless $t_c \in [1,2]$.

Lemma 2.4 will be proved in Section 3. Substituting (2.39) into (2.36) gives rise to two terms. Let us first estimate the contribution from the term containing $O_l(\frac{1}{|\eta|})$.

We bound it by

$$\iint \left| \hat{f}(\xi)\hat{g}(\eta)\hat{h}(\xi+\eta)\frac{1}{|\eta|} \right| \,\mathrm{d}\xi\,\mathrm{d}\eta \lesssim_l 2^{-k} \iint \left| \hat{f}(\xi)\hat{g}(\eta)\hat{h}(\xi+\eta) \right| \,\mathrm{d}\xi\,\mathrm{d}\eta. \tag{2.44}$$

By Hölder's inequality, the last term can be bounded by

$$2^{-k/2}\|f\|_2\|g\|_2\|h\|_2. \tag{2.45}$$

So far, we have managed to control the contribution from the 2nd term on the right-hand side of (2.39).

Now, we turn to the 1st term on the right-hand side of (2.39). The corresponding term we need to handle is

$$\iint \hat{f}(\xi)\hat{g}(\eta)\hat{h}(\xi+\eta)\frac{a(\xi,\eta)}{\sqrt{\eta}}e^{i\Psi(\xi,\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta. \tag{2.46}$$

We apply a change of variables $\xi \to 2^k\xi, \eta \to 2^k\eta$. We also rename $f, g, h$ for simplicity. It suffices to prove

$$\left| \iint f(\xi)g(\eta)h(\xi+\eta)a(\xi,\eta)e^{i2^k\Psi(\xi,\eta)}\,\mathrm{d}\xi\,\mathrm{d}\eta \right| \lesssim 2^{-\gamma k}\|f\|_2\|g\|_2\|h\|_2, \tag{2.47}$$

for every function $a : \mathbb{R}^2 \to \mathbb{R}$ with $\|a\|_{C^4} \lesssim 1$, and for functions $f$ supported on $\pm[1+\Delta_1 2^{-\gamma_0 l}, 1+(\Delta_1+2)2^{-\gamma_0 l}]$ and $g$ supported on $\pm[2^{-b_1}+\Delta_2 2^{-b_1-\gamma_0 l}, 2^{-b_1}+(\Delta_2+2)2^{-b_1-\gamma_0 l}]$. Here, $\Delta_1$ and $\Delta_2$ are two positive integers that are smaller than $2^{\gamma_0 l}$. Moreover, they are chosen such that (2.29) holds for every $\xi \in \mathrm{supp}(f)$ and $\eta \in \mathrm{supp}(g)$.

**Claim 2.5.** There exist integers $C_P$, $C'_P$ depending only on $P$ and intervals $J_1, \ldots, J_{C_P} \subset \mathbb{R}$ of length $2^{-\gamma k/C'_P}$, such that whenever $\xi/\eta \notin J_\iota$ for any $\iota$, we have

$$|\partial_\xi \partial_\eta(\partial_\xi - \partial_\eta)\Psi| \gtrsim 2^{-\gamma k}. \tag{2.48}$$

The implicit constant is allowed to depend on $P$ and can be taken to be polynomial in $2^l$.

The proof of the claim is postponed to the end of this section. Let $\widetilde{a} : \mathbb{R} \to \mathbb{R}$ be a smooth bump function taking value one on each $2J_\iota$ such that $\|\widetilde{a}\|_{C^4} \lesssim 2^{4\gamma k}$. To prove (2.47), we will decompose $a(\xi,\eta) = a(\xi,\eta)\widetilde{a}(\xi/\eta) + a(\xi,\eta)(1-\widetilde{a}(\xi/\eta))$ and control the two resulting terms separately. For the former term, the oscillation from $e^{i2^k\Psi}$ no longer plays any role, and we simply put the absolute value sign inside the integral and obtain

$$\iint |f(\xi)g(\eta)h(\xi+\eta)a(\xi,\eta)\widetilde{a}(\xi/\eta)|\,\mathrm{d}\xi\,\mathrm{d}\eta. \tag{2.49}$$

By Cauchy–Schwarz, this can be easily bounded by $2^{-\gamma k/C'_P}\|f\|_2\|g\|_2\|h\|_2$. To control the latter term, it suffices to prove the following lemma.

**Lemma 2.6.** For every small positive $\gamma > 0$, every function $a : \mathbb{R}^2 \to \mathbb{R}$ supported on $[-1,1]^2$ with $\|a\|_{C^4} \le 2^{\gamma k}$, every $\Psi : \mathbb{R}^2 \to \mathbb{R}$ with

$$|\partial_\xi \partial_\eta (\partial_\xi - \partial_\eta)\Psi| \gtrsim 2^{-\gamma k} \text{ and } \|\Psi\|_{C^4} \lesssim 1, \tag{2.50}$$

we have

$$\left| \iint f(\xi)g(\eta)h(\xi+\eta)a(\xi,\eta)e^{i2^k\Psi(\xi,\eta)}\, \mathrm{d}\xi\, \mathrm{d}\eta \right| \lesssim 2^{-\gamma k}\|f\|_2\|g\|_2\|h\|_2. \tag{2.51}$$

Here, taking $\gamma = 10^{-5}$ is more than enough.

**Proof of Lemma 2.6.** This lemma is essentially due to Li [12]. Here, we need to keep track of the dependence on norms of $a$, on its support, and so on. Oscillatory integrals of the form (2.51) have also been extensively studied in Xiao [16] and Gressman and Xiao [6].

We start the proof. By applying the triangle inequality, it suffices to prove (2.51) with a better gain $2^{-3\gamma k}$ in place of $2^{-\gamma k}$, for every function $g$ supported on an interval of length $2^{-2\gamma k}$. By a change of variable and by applying Cauchy–Schwarz, it is enough to prove

$$\left\| \int f(\xi-\eta)g(\eta)a(\xi-\eta,\eta)e^{i2^k\Psi(\xi-\eta,\eta)}\, \mathrm{d}\eta \right\|_{L^2_\xi}^2 \lesssim 2^{-6\gamma k}\|f\|_2^2\|g\|_2^2. \tag{2.52}$$

We expand the square on the left-hand side. After a change of variable, we obtain

$$\int_{|\tau|\le 2^{-\gamma k}} \iint_{\mathbb{R}^2} e^{i2^k[\Psi(\xi,\eta)-\Psi(\xi-\tau,\eta+\tau)]} F_\tau(\xi)G_\tau(\eta)a'_\tau(\xi,\eta)\, \mathrm{d}\xi\, \mathrm{d}\eta\, \mathrm{d}\tau, \tag{2.53}$$

for some new compactly supported amplitude $a'_\tau$. Moreover, $F_\tau(\cdot) := f(\cdot)\bar{f}(\cdot-\tau)$ and $G_\tau(\cdot) := g(\cdot)\bar{g}(\cdot+\tau)$. By the mean value theorem, it is easy to see that

$$\left|\partial_\xi\partial_\eta\left(\Psi(\xi,\eta)-\Psi(\xi-\tau,\eta+\tau)\right)\right| \gtrsim 2^{-\gamma k}|\tau|. \tag{2.54}$$

To proceed, we need the following lemma.

**Lemma 2.7.** For every small positive $\gamma > 0$, every function $a' : \mathbb{R}^2 \to \mathbb{R}$ supported on $[-1,1]^2$ with $\|a'\|_{C^4} \le 2^{\gamma k}$, every $\Xi : \mathbb{R}^2 \to \mathbb{R}$ with

$$\left|\partial_\xi \partial_\eta \Xi\right| \gtrsim 2^{-7\gamma k} \text{ and } \|\Xi\|_{C^4} \lesssim 1, \tag{2.55}$$

we have

$$\left|\iint F(\xi)G(\eta)a'(\xi,\eta)e^{i2^k \Xi(\xi,\eta)}\, \mathrm{d}\xi\, \mathrm{d}\eta\right| \lesssim 2^{-6\gamma k}\|F\|_2\|G\|_2. \tag{2.56}$$

Again, taking $\gamma = 10^{-5}$ is more than enough.

■

To control (2.53), we split the integral in $\tau$ into two parts:

$$\int_{|\tau|\le 2^{-6\gamma k}} + \int_{|\tau|\ge 2^{-6\gamma k}}. \tag{2.57}$$

Regarding the former term, we apply the triangle inequality and Cauchy–Schwarz to bound it by $2^{-6\gamma k}\|f\|_2^2\|g\|_2^2$. Regarding the latter term, we apply Lemma 2.7 and bound it by

$$2^{-6\gamma k}\int_{|\tau|\ge 2^{-6\gamma k}} \|F_\tau\|_2\|G_\tau\|_2\, \mathrm{d}\tau. \tag{2.58}$$

By applying Cauchy–Schwarz, this is bounded by $2^{-6\gamma k}\|f\|_2^2\|g\|_2^2$. This finishes the proof of Lemma 2.6.

**Proof of Lemma 2.7.** This lemma is essentially due to Hörmander [8]. By Cauchy–Schwarz, it suffices to prove

$$\left\|\int F(\xi)a'(\xi,\eta)e^{i2^k \Xi(\xi,\eta)}\, \mathrm{d}\xi\right\|_2^2 \lesssim 2^{-12\gamma k}\|F\|_2^2. \tag{2.59}$$

By the triangle inequality, it suffices to prove (2.59) with a better gain $2^{-30\gamma k}$ in place of $2^{-12\gamma k}$, for every function $F$ supported on an interval of length $2^{-8\gamma k}$. We expand the square on the left-hand side and obtain

$$\iint \sigma_k(\xi_1,\xi_2)F(\xi_1)\bar{F}(\xi_2)\, \mathrm{d}\xi_1\, \mathrm{d}\xi_2, \tag{2.60}$$

where

$$\sigma_k(\xi_1,\xi_2) := \int e^{i2^k(\Xi(\xi_1,\eta)-\Xi(\xi_2,\eta))}a'(\xi_1,\eta)\bar{a}'(\xi_2,\eta)\, \mathrm{d}\eta. \tag{2.61}$$

By the mean value theorem, we observe that

$$\left|\partial_\eta(\Xi(\xi_1,\eta)-\Xi(\xi_2,\eta))\right| \gtrsim 2^{-7\gamma k}|\xi_1-\xi_2|. \tag{2.62}$$

By applying integration by parts twice, we obtain

$$|\sigma_k(\xi_1,\xi_2)| \lesssim \min\{2^{2\gamma k}, 2^{-2k+50\gamma k}|\xi_1-\xi_2|^{-2}\}. \tag{2.63}$$

By Schur's test, this gives us the desired bound if we choose $\gamma$ small enough. This finishes the proof of Lemma 2.7. ■

**Proof of Claim 2.5.** Before properly embarking on the proof of this claim, we will summarize the proof. Hypothetically, if $\partial_\xi\partial_\eta(\partial_\xi-\partial_\eta)\Psi(\xi,\eta)$ were a nonzero polynomial function of $\frac{\xi}{\eta}$, Claim 2.5 would be trivial to prove by selecting the intervals $J_\iota$ to be neighborhoods of the zeros of $\partial_\xi\partial_\eta(\partial_\xi-\partial_\eta)\Psi(\xi,\eta)$. Unfortunately, $\partial_\xi\partial_\eta(\partial_\xi-\partial_\eta)\Psi(\xi,\eta)$ will not typically be a polynomial function of $\frac{\xi}{\eta}$. Nonetheless, we are able to compare $\partial_\xi\partial_\eta(\partial_\xi-\partial_\eta)\Psi(\xi,\eta)$ to such a polynomial after making a suitable change of variables.

Recall that $t_c(\xi,\eta)$ is defined via

$$\xi+a_1\eta+\eta Q'(2^{-l}t_c)=0. \tag{2.64}$$

Moreover,

$$\begin{aligned}\Psi(\xi,\eta) &= 2^{-l}(\xi+a_1\eta)t_c+\eta Q(2^{-l}t_c)\\ &= (\xi+a_1\eta)(Q')^{-1}\left(-\frac{\xi+a_1\eta}{\eta}\right)+\eta Q\left((Q')^{-1}\left(-\frac{\xi+a_1\eta}{\eta}\right)\right).\end{aligned} \tag{2.65}$$

Here, $(Q')^{-1}$ means the inverse of the derivative of $Q$. By a direct calculation presented in the appendix, we obtain

$$\begin{aligned}&\left|\partial_\xi\partial_\eta(\partial_\xi-\partial_\eta)\Psi(\xi,\eta)\right|\\ &\approx \left|(2\rho+1)\left(Q''\left((Q')^{-1}(-\rho-a_1)\right)\right)^2+(\rho^2+\rho)Q'''\left((Q')^{-1}(-\rho-a_1)\right)\right|,\end{aligned} \tag{2.66}$$

where $\rho:=\xi/\eta$. This statement means that these two expressions are the same up to a multiplicative factor that is bounded between $2^{-\delta l}$ and $2^{\delta l}$ for some real number $\delta$. By changing $\rho$ to $-\rho-a_1$, it is equivalent to consider

$$\left|-(2\rho+2a_1-1)\left(Q''\left((Q')^{-1}(\rho)\right)\right)^2+(\rho+a_1)(\rho+a_1-1)Q'''\left((Q')^{-1}(\rho)\right)\right|. \tag{2.67}$$

Recall that $Q(t) = a_2 t^{\alpha_2} + R(t)$, where $a_2 \neq 0$ and $R(t)$ can be viewed as a remainder term compared with $a_2 t^{\alpha_2}$ when $t \approx 2^{-l}$. Denote $s := (Q')^{-1}(\rho)$. Then, (2.67) becomes

$$\left| -(2Q'(s) + 2a_1 - 1)\left(Q''(s)\right)^2 + (Q'(s) + a_1)(Q'(s) + a_1 - 1)Q'''(s) \right|. \tag{2.68}$$

The highest-order term in the last display is given by

$$a_n^3 \alpha_n^3 s^{3\alpha_n - 5}\left(-2(\alpha_n - 1)^2 + (\alpha_n - 1)(\alpha_n - 2)\right). \tag{2.69}$$

Notice that the coefficient does not vanish. Therefore, we observe that (2.68) is the absolute value of a nonzero polynomial of $s$. By the observation from the summary of the proof, we have that there exist intervals $\tilde{J}_\iota$ such that $\partial_\xi \partial_\eta (\partial_\xi - \partial_\eta)\Psi(\xi, \eta)$ satisfies the desired inequality if $s \notin \tilde{J}_\iota$ for any $\iota$. But the transformation sending $\rho$ to $s$ is an invertible transformation with derivative bounded above and below polynomials in $2^{-l}$ as seen in the appendix, so the set of points $\rho$ such that $s \in \tilde{J}_\iota$ satisfy the conditions in Claim 2.5, provided that $\gamma$ is small enough. ■

## 3 Stationary Phase Principle: Proof of Lemma 2.4

Our goal in this section is to prove an asymptotic formula for

$$\int_{\mathbb{R}} e^{i2^{-l}t\xi + iP(2^{-l}t)\eta}\tau_0(t)\,\mathrm{d}t = 2^l \int_{\mathbb{R}} e^{it\xi + iP(t)\eta}\tau_0(2^l t)\,\mathrm{d}t. \tag{3.1}$$

We follow the proof of Proposition 3 on Stein [18, p. 334]. Define

$$\Phi_{\xi,\eta}(t) := t\xi + P(t)\eta. \tag{3.2}$$

Recall some notation

$$P(t) = a_1 t + Q(t) = a_1 t + a_2 t^{\alpha_2} + R(t). \tag{3.3}$$

Let $\bar{t}_c$. be the critical point of the phase function, that is,

$$\xi + P'(\bar{t}_c)\eta = \xi + (a_1 + Q'(\bar{t}_c))\eta = 0. \tag{3.4}$$

By a rescaling of the observation preceding Lemma 2.4, we have that $\bar{t}_c$ lies in $(2^{-l-2}, 2^{-l+2})$. We expand the phase function about $\bar{t}_c$:

$$\Phi_{\xi,\eta}(t) = \Phi_{\xi,\eta}(\bar{t}_c) + \frac{1}{2}Q''(\bar{t}_c)\eta(t - \bar{t}_c)^2 + O_l(|t - \bar{t}_c|^3)\cdot\eta. \tag{3.5}$$

Here,

$$O_l(|t-\bar{t}_c|^3) \leq C_P 2^l |Q''(\bar{t}_c)||t-\bar{t}_c|^3, \tag{3.6}$$

where $C_P$ is a large constant depending only on $P$. Let $\vartheta$ be a nonnegative even smooth function supported on $[-2,2]$, constant on $[-1,1]$ and monotone on $[1,2]$. We normalize it such that $\widehat{\vartheta}(0)=1$ and denote $\vartheta_\ell(x) := 2^\ell \vartheta(2^\ell x)$. We write

$$\begin{aligned} \int_{\mathbb{R}} e^{it\xi+iP(t)\eta}\tau_0(2^l t)\,\mathrm{d}t = & \int_{\mathbb{R}} e^{it\xi+iP(t)\eta}\tau_0(2^l t)\vartheta(2^{l+10C_P}(t-\bar{t}_c))\,\mathrm{d}t \\ & + \int_{\mathbb{R}} e^{it\xi+iP(t)\eta}\tau_0(2^l t)\left(1-\vartheta(2^{l+10C_P}(t-\bar{t}_c))\right)\,\mathrm{d}t =: I + II. \end{aligned} \tag{3.7}$$

The phase function in term $II$ does not admit any critical point. Hence, by integration by parts, we obtain

$$|II| \leq 2^{\gamma_0 l}\frac{1}{|\eta|}. \tag{3.8}$$

For term $I$, we write it as

$$e^{i\Phi_{\xi,\eta}(\bar{t}_c)}\int_{\mathbb{R}} e^{i\Phi_{\xi,\eta}(t)-i\Phi_{\xi,\eta}(\bar{t}_c)}\widetilde{\vartheta}_l(t)\,\mathrm{d}t, \tag{3.9}$$

where

$$\widetilde{\vartheta}_l(t) := \tau_0(2^l t)\vartheta(2^{l+10C_P}(t-\bar{t}_c)). \tag{3.10}$$

The support of the function $\widetilde{\vartheta}_l$ is chosen to be so small such that the change of variable

$$(t-\bar{t}_c)^2 + \frac{2}{Q''(\bar{t}_c)}\cdot O_l(|t-\bar{t}_c|^3) \to y^2 \tag{3.11}$$

becomes valid. Under this change of variable, (3.9) turns to

$$e^{i\Phi_{\xi,\eta}(\bar{t}_c)}\int_{\mathbb{R}} e^{i\frac{1}{2}Q''(\bar{t}_c)\eta y^2}\vartheta_l'(y)\,\mathrm{d}y, \tag{3.12}$$

for some new smooth truncation function $\vartheta_l'$. We split the last expression into three terms:

$$e^{i\Phi_{\xi,\eta}(\bar{t}_c)}\int_{\mathbb{R}} e^{i\frac{1}{2}Q''(\bar{t}_c)\eta y^2}e^{-y^2}(e^{y^2}\vartheta_l'(y)-\vartheta_l'(0))\vartheta_l''(y)\,\mathrm{d}y$$
$$+e^{i\Phi_{\xi,\eta}(\bar{t}_c)}\int_{\mathbb{R}} e^{i\frac{1}{2}Q''(\bar{t}_c)\eta y^2}e^{-y^2}\vartheta_l'(0)(\vartheta_l''(y)-1)\,\mathrm{d}y \tag{3.13}$$
$$+e^{i\Phi_{\xi,\eta}(\bar{t}_c)}\int_{\mathbb{R}} e^{i\frac{1}{2}Q''(\bar{t}_c)\eta y^2}e^{-y^2}\vartheta_l'(0)\,\mathrm{d}y,$$

where $\vartheta_l''$ is a compactly supported smooth function and is 1 on the support of $\vartheta_l'$. These three terms will be called $I_1, I_2$, and $I_3$ and will be handled separately.

By the triangle inequality and an integration by parts argument, it is not difficult to see that

$$|I_1|+|I_2|\le 2^{\gamma_0 l}|\eta|^{-1}. \tag{3.14}$$

In the end, one just needs to observe that

$$\int_{\mathbb{R}} e^{i\lambda t^2}e^{-t^2}\,\mathrm{d}t = e_0\lambda^{-1/2}+O(\lambda^{-3/2}) \tag{3.15}$$

for some universal constant $e_0$. See Stein [18, Equation (9) on p. 335]. This finishes the proof of Lemma 2.4.

## 4 Positivity of the Double Integral: Proof of Theorem 1.3

In this section, we prove Theorem 1.3. We follow the idea of Łaba and Pramanik [11] and decompose

$$\mu=\mu_1+\mu_2, \tag{4.1}$$

with

$$\mu_1(x)\le A\cdot 2^{6B}C_1, \tag{4.2}$$

where $A$ is a large absolute constant. Here, $\mu_1$ is obtained by convolving $\mu$ with a Fejér kernel. See Łaba and Pramanik [11, p. 442]. Here, we make a remark that this is the only place where one applies the assumption $(A)$ in (1.1). Also, in their decomposition, it is possible to choose $\mu_1$ so that

$$\mu_1\ge 0 \text{ and } \int_0^1\mu_1(x)\,\mathrm{d}x=1. \tag{4.3}$$

Moreover, we have $\widehat{\mu_2}(0) = 0$, and

$$\widehat{\mu_2}(n) = \min\left(1, \frac{|n|}{2N+1}\right)\hat{\mu}(n), \tag{4.4}$$

where

$$N = C_2^{-1} e^{\frac{1}{1-\alpha}}. \tag{4.5}$$

**Lemma 4.1.** There exists $l_0 \in \mathbb{N}$ and $c_0 > 0$ depending only on $C_1, C_2, B$, $\beta$ and the polynomial $P$ such that

$$\iint \mu_1(x)\mu_1(x+t)\mu_1(x+P(t))\tau_{l_0}(t)\,\mathrm{d}t\,\mathrm{d}x \geq c_0. \tag{4.6}$$

The proof of Lemma 4.1 is based on Bourgain's energy pigeonholing argument [3] and the Sobolev improving estimate in Proposition 1.2. We postpone its proof to the next section.

After finding $l_0$ and $c_0$, we will pick $\alpha$ to be sufficiently close to one and prove that

$$\left|\iint \mu_{i_1}(x)\mu_{i_2}(x+t)\mu_{i_3}(x+P(t))\tau_{l_0}(t)\,\mathrm{d}t\,\mathrm{d}x\right| \leq c_0/8, \tag{4.7}$$

when $(i_1, i_2, i_3) \neq (1,1,1)$. For the sake of simplicity, let us assume that we are working with $(i_1, i_2, i_3) = (1,1,2)$. The proofs of the other cases are similar. In the previous sections, we proved that

$$\left|\iint \mu_1(x)\mu_1(x+t)\mu_2(x+P(t))\tau_{l_0}(t)\,\mathrm{d}t\,\mathrm{d}x\right| \leq C_{l_0}\|\mu_1\|_{H^{-s_0}}^2\|\mu_2\|_{H^{-s_0}}, \tag{4.8}$$

for some $s_0 > 0$ depending only on the polynomial $P$. By the definition of $\mu_1$ and the assumption on $\mu$, we have

$$\|\mu_1\|_{H^{-s_0}}^2 \leq \|\mu\|_{H^{-s_0}}^2 \leq C_2^2(1-\alpha)^{-2B}\sum_{k\geq 1}|k|^{-\beta}|k|^{-2s_0}. \tag{4.9}$$

Next, we turn to the term $\|\mu_2\|_{H^{-s_0}}$.

$$\begin{aligned}\|\mu_2\|_{H^{-s_0}}^2 &\leq C_2^2(1-\alpha)^{-2B}\left(\sum_{1\leq k\leq 2N}\frac{k^2}{(2N+1)^2}k^{-\beta-2s_0} + \sum_{k>2N}k^{-\beta-2s_0}\right)\\ &\lesssim C_2^2(1-\alpha)^{-2B}\left(\frac{N^{3-\beta-2s_0}}{N^2} + N^{1-\beta-2s_0}\right) \leq C_2^2(1-\alpha)^{-2B}N^{1-\beta-2s_0}.\end{aligned} \tag{4.10}$$

Combined with (4.8), we obtain

$$\left|\iint \mu_1(x)\mu_1(x+t)\mu_2(x+P(t))\tau_{l_0}(t)\,\mathrm{d}t\,\mathrm{d}x\right| \le C_{l_0}C_2^{10}(1-\alpha)^{-10B}N^{(1-\beta-2s_0)/2}. \tag{4.11}$$

Recall (4.5). If we choose $\alpha$ close enough to one, depending on all the other parameters, we will be able to conclude (4.7).

## 5 Proof of Lemma 4.1

Before we start the proof of Lemma 4.1, we state a preliminary lemma. Recall the definition of $\vartheta$ in Section 3.

**Lemma 5.1** (Bourgain [3]). For a nonnegative function $f$ supported on $[0,1]$ and $k, l \in \mathbb{N}$, we have

$$\int_0^1 f(f*\vartheta_k)(f*\vartheta_\ell) \ge c_0\left(\int_0^1 f\right)^3$$

for some constant $c_0 > 0$ depending only on the choice of $\vartheta$.

The proof of this lemma was omitted in [3]. For a proof, we refer to [5].

In this section, we will use $f$ to stand for $\mu_1$. Hence, $f$ is a function satisfying

$$\int f = 1 \text{ and } 0 \le f \le A\cdot 2^{6B}C_1 =: M. \tag{5.1}$$

For simplicity, we assume $||\tau_0||_1 = 1$ and change the notation a bit by taking

$$\tau_l(t) = 2^l\tau_0(2^l t) \text{ instead of } \tau_l(t) = \tau_0(2^l t). \tag{5.2}$$

We also need to show that $l_0$ can be bounded from above by a number that depends only on $C_1$, $C_2$, $B$, $\beta$, and $P$. Denote

$$\Lambda_l = \iint f(x)f(x+t)f(x+P(t))\tau_l(t)\,\mathrm{d}t\,\mathrm{d}x. \tag{5.3}$$

For $\ell', \ell, \ell'' \in \mathbb{N}$ with $1 \le \ell' \le \ell \le \ell''$, we have

$$\begin{aligned}\Lambda_l &= \int_0^1\int_0^1 f(x)f(x+t)f(x+P(t))\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t\\ &= I_1 + I_2 + I_3,\end{aligned}$$

where

$$I_1 = \int_0^1 \int_0^1 f(x)f(x+t)(f*\vartheta_{\ell'})(x+P(t))\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t,$$
$$I_2 = \int_0^1 \int_0^1 f(x)f(x+t)(f*\vartheta_{\ell''}-f*\vartheta_{\ell'})(x+P(t))\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t,$$
$$I_3 = \int_0^1 \int_0^1 f(x)f(x+t)(f-f*\vartheta_{\ell''})(x+P(t))\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t.$$

We analyze each of the terms separately. Splitting $f-f*\vartheta_{\ell''}$ into Littlewood–Paley pieces and applying Lemmas 2.2 and 2.1, it follows that for some $\sigma>0$, we have

$$|I_3| \lesssim_M 2^{\gamma_0\ell-\sigma\ell''}\|f\|_{L^2(\mathbb{R})}^2 \le 2^{-100}c_0,$$

where the last inequality holds provided that $\ell'' > \frac{\gamma_0}{\sigma}\ell + C(M)$, which will hold for sufficiently large $\ell$ provided that $\ell'' > 2\frac{\gamma_0}{\sigma}\ell$. Here, $c_0$ is the constant from Lemma 5.1 and $C(M)$ is a quantity depending on $M$, $\sigma$, and $c_0$ but not on $\ell$ or $\ell'$. We have also applied the pointwise bound 5.1, resulting a square of $\|f\|_2$ instead of a cubic power.

To estimate $I_2$, we apply the Cauchy–Schwarz inequality in $x$, which yields

$$\begin{aligned}|I_2| &\le \int_0^1 \|f(x)f(x+t)\|_{L^2_x}\|(f*\vartheta_{\ell''}-f*\vartheta_{\ell'})(x+P(t))\|_{L^2_x}\tau_\ell(t)\,\mathrm{d}t\\ &\lesssim_M \|f*\vartheta_{\ell''}-f*\vartheta_{\ell'}\|_2.\end{aligned}$$

Passing to the last line, we bounded the $L^\infty$ norm of $f$ by $M$ and the $L^1$ norm of $\tau_\ell$ by one.

To estimate $I_1$, we compare it with

$$\begin{aligned}I_4 &= \int_0^1 \int_0^1 f(x)f(x+t)(f*\vartheta_{\ell'})(x)\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t\\ &= \int_0^1 f(x)(f*\vartheta_{\ell'})(x)(f*\tau_\ell)(x)\,\mathrm{d}x.\end{aligned}$$

Consider the difference

$$I_4 - I_1 = \int_0^1 \int_0^1 f(x)f(x+t)\left((f*\vartheta_{\ell'})(x)-(f*\vartheta_{\ell'})(x+P(t))\right)\tau_\ell(t)\,\mathrm{d}x\,\mathrm{d}t.$$

By the mean value theorem, we obtain

$$|(f*\vartheta_{\ell'})(x)-(f*\vartheta_{\ell'})(x+P(t))| \le 2^{\ell'}\|f*(\vartheta_{\ell'})'\|_\infty |P(t)| \le M2^{\ell'-\ell+1},$$

whenever $t$ is in the support of $\tau_\ell$. If $\ell'$ is sufficiently large depending on $M$ and $c_0$, then choosing $\ell > 2\ell'$ gives

$$|I_4 - I_1| \leq 2^{-100} c_0.$$

We return to analyzing the term $I_4$, which we write as

$$I_4 = \left( \int_0^1 f(x)(f * \vartheta_{\ell'})(x) \left( (f * \tau_\ell)(x) - (f * \vartheta_{\ell'})(x) \right) \, \mathrm{d}x \right) \tag{5.4}$$

$$+ \left( \int_0^1 f(x)(f * \vartheta_{\ell'})(x)(f * \vartheta_{\ell'})(x) \, \mathrm{d}x \right). \tag{5.5}$$

By Lemma 5.1, the term (5.5) is bounded from below by $c_0$. For (5.4), we use the triangle inequality and Young's convolution inequality to estimate

$$\|f * \tau_\ell - f * \vartheta_{\ell'}\|_2$$

$$\lesssim_M \|(f * \tau_\ell * \vartheta_{\ell''}) - (f * \vartheta_{\ell'} * \tau_\ell)\|_2 \tag{5.6}$$

$$+ \|\tau_\ell - (\tau_\ell * \vartheta_{\ell''})\|_1 \tag{5.7}$$

$$+ \|\vartheta_{\ell'} - (\vartheta_{\ell'} * \tau_\ell)\|_1. \tag{5.8}$$

By another application of Young's convolution inequality in (5.6), we bound the last display by

$$\lesssim_M \|(f * \vartheta_{\ell''}) - (f * \vartheta_{\ell'})\|_2 + \|\tau_\ell - (\tau_\ell * \vartheta_{\ell''})\|_1 + \|\vartheta_{\ell'} - (\vartheta_{\ell'} * \tau_\ell)\|_1.$$

By the mean value theorem, the 2nd and 3rd terms are bounded from above by $2^{-100} c_0$ provided $\ell'$ is chosen large enough and that $\ell > 2\ell'$ and $\ell'' > 2\ell$. This in turn bounds (5.4) from above by

$$\|f * \vartheta_{\ell''} - f * \vartheta_{\ell'}\|_2 + 2^{-99} c_0.$$

From the estimates for the terms $I_1, I_2, I_3, I_4$, and $I_4 - I_1$, we obtain

$$c_0 \leq \Lambda_l + C_M \|f * \vartheta_{\ell'} - f * \vartheta_{\ell''}\|_2 + 2^{-90} c_0.$$

Here, $C_M$ is a large constant that depends only on $M$. For instance, it suffices to take $C_M = M^{10}$. Therefore, we have either $\Lambda_l > 2^{-10} c_0$ or

$$\|f * \vartheta_{\ell'} - f * \vartheta_{\ell''}\|_2 > 2^{-10} C_M^{-1} c_0.$$

By the preceding discussion, we can construct a sequence $\{\ell_0 < \ell_1 < \cdots < \ell_k < \cdots\} \subseteq \mathbb{N}$, which is independent of $f$ and satisfies $\ell_{k+1} \leq C\ell_k$ for some sufficiently large constant $C$ such that for each $k$ either

$$\Lambda_{\ell_k} > 2^{-10}c_0$$

or

$$\|f * \vartheta_{\ell_k} - f * \vartheta_{\ell_{k+1}}\|_2 > 2^{-10}c_0 C_M^{-1}. \tag{5.9}$$

Observe that for any $K \geq 0$, one has

$$\sum_{k=0}^{K} \left(\|f * \vartheta_{\ell_k} - f * \vartheta_{\ell_{k+1}}\|_2^2\right) \leq C_0\|f\|_2^2 \leq C_0 M^2 \tag{5.10}$$

with $C_0$ independent of $K$ and $f$. Let us fix $K > C_0 2^{100} c_0^{-2} C_M^2 M^4$. If (5.9) holds for all $0 < k \leq K$, then (5.10) yields $K \leq C_0 2^{100} c_0^{-2} C_M^2 M^4$, which is a contradiction. Thus, for some $0 \leq k \leq K$, we necessarily have $\Lambda_{\ell_k} > 2^{-10}c_0$. Together with $\ell_{k+1} \leq C\ell_k$, this gives a lower estimate on $\Lambda_{\ell_k}$, as claimed in Lemma 4.1.

## 6 Existence of Polynomial Patterns: Proof of Theorem 1.1

Let $\mu$ be as in Theorem 1.1. Let $l_0$ be as in Theorem 1.3. Theorem 1.1 follows if we are able to construct a Borel measure $\nu$ on $[0,1] \times [0,1]$ such that

$$\nu([0,1] \times [0,1]) > 0 \tag{6.1}$$

and

$$\nu \text{ is supported on} X = \{(x,y) \in [0,1]^2 : x, y, x + P(y-x) \in E,\ 2^{-l_0} \leq y - x \leq 2^{-l_0+1}\}. \tag{6.2}$$

For $\epsilon > 0$, define

$$\mu_\epsilon := \mu * \vartheta_\epsilon, \tag{6.3}$$

where $\vartheta_\epsilon(x) = \epsilon^{-1}\vartheta(x/\epsilon)$. A standard argument shows that

$$\mu_\epsilon \to \mu \text{ in } H^{-s_0} \text{ as } \epsilon \to 0. \tag{6.4}$$

We define a linear functional $\nu$ acting on functions $f : [0,1]^2 \to \mathbb{R}$ by

$$\langle \nu, f\rangle := \lim_{\epsilon\to 0} \iint f(x,y)\mu_\epsilon(x+P(y-x))\tau_{l_0}(y-x)\,\mathrm{d}\mu_\epsilon(x)\,\mathrm{d}\mu_\epsilon(y). \tag{6.5}$$

The following lemma holds.

**Lemma 6.1.** The limit in (6.5) exists for every continuous function $f$. Moreover,

$$|\langle \nu, f\rangle| \le C\|f\|_\infty, \tag{6.6}$$

where $C$ is independent of $f$.

**Proof.** of Lemma 6.1 For every $\epsilon > 0$, the following inequality holds:

$$\begin{aligned} &\left|\iint f(x,y)\mu_\epsilon(x+P(y-x))\tau_{l_0}(y-x)\,\mathrm{d}\mu_\epsilon(x)\,\mathrm{d}\mu_\epsilon(y)\right| \\ &\le \|f\|_\infty \iint \mu_\epsilon(x+P(y-x))\tau_{l_0}(y-x)\,\mathrm{d}\mu_\epsilon(x)\,\mathrm{d}\mu_\epsilon(y). \end{aligned} \tag{6.7}$$

By the Sobolev improving estimate in Proposition 1.2, this can be bounded by

$$2^{\gamma_0 l_0}\|\mu_\epsilon\|^2_{H^{-s_0}}\|\mu_\epsilon\|_{H^{-s_0}} \le 2^{3+\gamma_0 l_0}\|\mu\|^3_{H^{-s_0}}. \tag{6.8}$$

Recall that $1 - s_0 < \beta$. Under this assumption, we know $\|\mu\|_{H^{-s_0}}$ is finite. This proves (6.6) if the limit (6.5) exists.

It remains to prove the existence of the limit (6.5). By density, it suffices to prove that the limit exists for every smooth function $f$ whose Fourier series consists of only finitely many terms. Hence, it suffices to prove that the limit

$$\lim_{\epsilon\to 0} \iint e^{iN_1x+iN_2y}\mu_\epsilon(x+P(y-x))\tau_{l_0}(y-x)\,\mathrm{d}\mu_\epsilon(x)\,\mathrm{d}\mu_\epsilon(y) \tag{6.9}$$

exists for given $N_1, N_2 \in \mathbb{N}$. By Proposition 1.2,

$$\begin{aligned} &\left|\iint e^{iN_1x+iN_2y}(\mu_{\epsilon_1}-\mu_{\epsilon_2})(x+P(y-x))\tau_{l_0}(y-x)\,\mathrm{d}\mu_\epsilon(x)\,\mathrm{d}\mu_\epsilon(y)\right| \\ &\le 2^{\gamma_0 l_0}\|\mu_{\epsilon_1}-\mu_{\epsilon_2}\|_{H^{-s_0}}\|\mu'\|_{H^{-s_0}}\|\mu''\|_{H^{-s_0}}, \end{aligned} \tag{6.10}$$

where

$$d\mu'(x) = e^{iN_1 x} d\mu_\epsilon(x) \text{ and } d\mu''(x) = e^{iN_2 x} d\mu_\epsilon(x). \tag{6.11}$$

The right-hand side of (6.10) can be further bounded by

$$C_{N_1,N_2} 2^{\gamma_0 l_0} \|\mu_{\epsilon_1} - \mu_{\epsilon_2}\|_{H^{-s_0}} \|\mu\|^2_{H^{-s_0}}. \tag{6.12}$$

That the limit exists follows from the fact that $\|\mu_{\epsilon_1} - \mu_{\epsilon_2}\|_{H^{-s_0}}$ can be made arbitrarily small when $\epsilon_1$ and $\epsilon_2$ are chosen small enough. This finishes the proof of Lemma 6.1. ■

After obtaining Lemma 6.1, we apply the Riesz representation theorem and obtain a nonnegative measure $\nu$ defined by (6.5). It remains to prove that $\nu$ satisfies the desired properties (6.1) and (6.2).

To prove (6.1), we write

$$\begin{aligned} \langle \nu, 1\rangle &= \lim_{\epsilon\to 0} \iint \mu_\epsilon(x+P(t))\tau_{l_0}(t)\, \mathrm{d}\mu_\epsilon(x)\, \mathrm{d}\mu_\epsilon(x+t) \\ &= \iint_{\mathbb{R}^2} \widehat{\mu}(\xi)\widehat{\mu}(\eta)\left[\int_{\mathbb{R}} e^{it\xi + iP(t)\eta}\tau_{l_0}(t)\,\mathrm{d}t\right]\widehat{\mu}(\xi+\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta. \end{aligned} \tag{6.13}$$

From Theorem 1.3, it follows that $\langle \nu, 1\rangle \geq c_0 > 0$. This proves (6.1).

Finally, we prove (6.2). Let us introduce

$$\widetilde{X} := \{(x,y) \in [0,1]^2 : x, y, x+P(y-x) \in E\}. \tag{6.14}$$

By the definition of the measure $\nu$, it is enough to prove that $\nu$ is supported on $\widetilde{X}$. Let $f$ be a continuous function with $\mathrm{supp}(f)$ disjoint from $\widetilde{X}$. We need to prove that $\langle \nu, f\rangle = 0$. Since $E$ is closed, $\widetilde{X}$ is also closed. Moreover, $\mathrm{dist}(\mathrm{supp}(f), \widetilde{X}) > 0$. Using a partition of unity, we are able to write $f$ as a finite sum $\sum f_j$, where for each $j$, the function $f_j$ is continuous and satisfies at least one of the following:

$$\begin{aligned} &\mathrm{dist}\,(\mathrm{supp}(f_j), E\times[0,1]) > 0, \\ &\mathrm{dist}\,(\mathrm{supp}(f_j), [0,1]\times E\}) > 0, \\ &\mathrm{dist}\,\left(\left\{x+P(y-x) : (x,y)\in \mathrm{supp}(f_j)\right\}, E\right) > 0. \end{aligned} \tag{6.15}$$

We will prove that $\langle \nu, f_j \rangle = 0$ for every $j$. If $f_j$ satisfies either the 1st or the 2nd condition in (6.15), then the integral in (6.5) is 0 for every $\epsilon$ small enough. If $f_j$ satisfies the 3rd condition in (6.15), then the support of $f_j$ is a positive distance from the support of $\mu_\epsilon(x+P(y-x)$ for sufficiently small $\epsilon$, so the integral is again equal to 0 if $\epsilon$ is sufficiently small. This finishes the proof of (6.2).

## A Appendix

We compute the derivative $(\partial_{\xi,\xi,\eta} - \partial_{\xi,\eta,\eta})\Psi(\xi,\eta)$ and establish the approximation (2.66) for $t_c \in [1,2)$.

Recall that we have the expression $P(t) = a_1 t + Q(t)$, where the lowest-degree term of $Q(t)$ is $a_2 t^{\alpha_2}$. The condition (2.4) on $l$ guarantees that $a_2 t^{\alpha_2}$ is the dominant term of $Q(t)$.

We will make the following claims:

(i) $|Q''(2^{-l}t)|$ is bounded above and below by polynomial functions of $2^{-l}$, for $t \in [1,2)$;

(ii) $Q'$ is monotone and hence invertible on $[2^{-l}, 2^{-l+1})$ and $|((Q')^{-1})'|$ is bounded from above and below by a polynomial in $2^l$.

The lowest-degree term of $Q''(t)$ is $\alpha_2(\alpha_2 - 1)a_2 t^{\alpha_2 - 2}$. By the condition (2.4) on $l$, we have that, for $t \in [1,2)$, $|Q''(2^{-l}t)|$ is bounded between $2^{-(\alpha_2-2)l-1}\alpha_2(\alpha_2 - 1)|a_2|$ and $2^{-(\alpha_2-2)(l-1)+1}(\alpha_2 - 1)\alpha_2|a_2|$ for $t \in [1,2)$. Both bounds are polynomial in $2^{-l}$, establishing (i). So, $Q'$ is monotone (and hence invertible) on $[2^{-l}, 2^{-l+1})$ and, by the inverse function theorem, we have that $|((Q')^{-1})'|$ is bounded from above and below by a polynomial in $2^l$. This establishes (ii).

Now, we are ready to prove the inequality. We will compute the derivative $\partial_\xi \partial_\eta (\partial_\xi - \partial_\eta)\Psi(\xi,\eta)$. We will start with by taking the derivative of (2.42) with respect to $\eta$. This derivative is

$$\partial_{\xi,\eta}\Psi(\xi,\eta) = \frac{\xi}{\eta^2}((Q')^{-1})'\left(-\frac{\xi + a_1\eta}{\eta}\right). \tag{A.1}$$

The 3rd derivative $\partial_{\xi,\xi,\eta}\Psi(\xi,\eta)$ is given by

$$\partial_{\xi,\xi,\eta}\Psi(\xi,\eta) = \frac{1}{\eta^2}((Q')^{-1})'\left(-\frac{\xi + a_1\eta}{\eta}\right) - \frac{\xi}{\eta^3}((Q')^{-1})''\left(-\frac{\xi + a_1\eta}{\eta}\right). \tag{A.2}$$

The 3rd derivative $\partial_{\xi,\eta,\eta}\Psi(\xi,\eta)$ is given by

$$\partial_{\xi,\eta,\eta}\Psi(\xi,\eta) = -\frac{2\xi}{\eta^3}((Q')^{-1})'\left(-\frac{\xi + a_1\eta}{\eta}\right) + \frac{\xi^2}{\eta^4}((Q')^{-1})''\left(-\frac{\xi + a_1\eta}{\eta}\right). \tag{A.3}$$

So, the difference is

$$\left(\frac{1}{\eta^2}+2\frac{\xi}{\eta^3}\right)((Q')^{-1})'\left(-\frac{\xi+a_1\eta}{\eta}\right)+\left(-\frac{\xi}{\eta^3}-\frac{\xi^2}{\eta^4}\right)((Q')^{-1})''\left(-\frac{\xi+a_1\eta}{\eta}\right). \tag{A.4}$$

We make the substitution $\rho=\frac{\xi}{\eta}$, yielding

$$\frac{1}{\eta^2}\Big[(1+2\rho)((Q')^{-1})'(-\rho-a_1)-(\rho^2+\rho)((Q')^{-1})''(-\rho+a_1)\Big]. \tag{A.5}$$

But we also have

$$((Q')^{-1})'=\frac{1}{Q''((Q')^{-1})} \tag{A.6}$$

and

$$((Q')^{-1})''=-\frac{Q'''((Q')^{-1})}{Q''((Q')^{-1})^2}((Q')^{-1})'. \tag{A.7}$$

Therefore, we get

$$\begin{aligned}&\left(\frac{\eta^2(Q''((Q')^{-1}(-\rho-a_1)))^2}{((Q')^{-1})'(-\rho-a_1)}\right)(\partial_{\xi,\xi,\eta}-\partial_{\xi,\eta,\eta})\Psi(\xi,\eta)\\ =&(1+2\rho)(Q''((Q')^{-1}(-\rho-a_1)))^2+(\rho^2+\rho)Q'''((Q')^{-1}(-\rho-a_1)).\end{aligned} \tag{A.8}$$

The factor $\frac{\eta^2(Q''((Q')^{-1}(-\rho-a_1)))^2}{((Q')^{-1})'(-\rho-a_1)}$ is bounded between $2^{-\delta l}$ and $2^{\delta l}$ for some $\delta$ by (i) and (ii), as well as the fact that $\eta\approx 1$. This proves the desired approximation (2.66).

**Funding**

This work was supported by the National Science Foundation [1803086 to R.F., 1800274 to S.G.]; the Chinese University of Hong Kong [4053295 to S.G.]; Natural Sciences and Engineering Research Council of Canada discovery grant; National Science Foundation [DMS-1440140].

**Acknowledgments**

The authors would like to thank the referees for reading the manuscript carefully and for providing a number of valuable suggestions.

## References

[1] Bluhm, C. "Random recursive construction of Salem sets." *Ark. Mat.* 34 (1996): 51–63.
[2] Bluhm, C. "On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets." *Ark. Mat.* 36 (1998): 307–16.
[3] Bourgain, J. "A nonlinear version of Roth's theorem for sets of positive density in the real line." *J. Anal. Math.* 50 (1988): 169–81.
[4] Chan, V., I. Łaba, and M. Pramanik. "Finite configurations in sparse sets." *J. Anal. Math.* 128 (2016): 289–335.

[5] Durcik, P., S. Guo, and J. Roos. "A polynomial Roth theorem on the real line." *Trans. Amer. Math. Soc.* 371, no. 10 (2019): 6973–93.

[6] Gressman, P. and L. Xiao. "Maximal decay inequalities for trilinear oscillatory integrals of convolution type." *J. Funct. Anal.* 271, no. 12 (2016): 3695–726.

[7] Henriot, K., I. Łaba, and M. Pramanik. "On polynomial configurations in fractal sets." *Anal. PDE* 9, no. 5 (2016): 1153–84.

[8] Hörmander, L. "Oscillatory integrals and multipliers on $FL^p$." *Ark. Mat* 11, no. 1 (1973): 1–11.

[9] Kahane, J. P. *Some Random Series of Functions*. Cambridge: Cambridge University Press, 1985.

[10] Kaufman, L. "On the theorem of Jarnik and Besicovitch." *Acta Arith.* 39 (1981): 265–7.

[11] Łaba, I. and M. Pramanik. "Arithmetic progressions in sets of fractional dimension." *Geom. Funct. Anal.* 19, no. 2 (2009): 429–56.

[12] Li, X. "Bilinear Hilbert transforms along curves I: the monomial case." *Anal. PDE* 6, no. 1 (2013): 197–220.

[13] Lie, V. "On the boundedness of the bilinear Hilbert transform along "non-flat" smooth curves." *Amer. J. Math.* 137, no. 2 (2015): 313–63.

[14] Mattila, P. *Geometry of Sets and Measures in Euclidean Space*. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 1995.

[15] Shmerkin, P. "Salem sets with no arithmetic progressions." *Int. Math. Res. Not. IMRN*, 2017, no. 7 (2017): 1929–41.

[16] Xiao, L. "Sharp estimates for trilinear oscillatory integrals and an algorithm of two-dimensional resolution of singularities." *Rev. Mat. Iberoam.* 33, no. 1 (2017): 67–116.

[17] Salem, R. "On singular monotonic functions whose spectrum has a given Hausdorff dimension." *Ark. Mat.* 1 (1950): 353–65.

[18] Stein, E. M. *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton: Princeton University Press, 1993.